CLUSTERS OF THE CZECH REGIONS AND 31 CHOSEN EUROPEAN COUNTRIES ACCORDING TO THE WAGE LEVEL, WAGE MODELS AND GENDER WAGE GAP BY THE CZECH REGIONS

DIANA BÍLKOVÁ

University of Economics, Prague, Faculty of Informatics and Statistics,
Sq. W. Churchill 1938/4, 130 67 Prague 3, Czech Republic
email: bilkova@vse.cz

Abstract
The main aim of this paper is the creation of the clusters of the Czech regions that are close as possible by the wage level in these regions and clusters of 31 chosen European countries according to the same criterion. The Czech regions were clustered into three or five clusters, while the European countries were clustered into three, five or seven clusters. The number of clusters has been selected based on the number of clustered objects and due to the clarity of the results obtained. Using the methods of cluster analysis, the clusters with similar wage level were created. Methods of the furthest neighbour and Euclidean distance metric were used. For the Czech regions, the model wage distribution was constructed to enable the comparison of wage development during recent seven years. Three-parametric lognormal curves present the basis of the theoretical wage distribution. The special attention was paid to the situation of wage level within European countries and its comparison with the wage level in the Czech Republic. The researched variable is the gross monthly (nominal) wage. Database comes from the official website of the Czech Statistical Office.

Key words: *clusters of the Czech regions by wage level, clusters of 31 chosen European countries by wage level, wage models in the Czech regions, gender wage gap by the Czech regions, prediction of gender wage gap by the Czech regions*
JEL Codes: *J31, D31, E24, C10*

1. Introduction

Population wage levels are an object of research for economists of all developed countries, since they reflect the living standards. Knowledge of wage distribution of the population and its comparison from various socio-economic and time-spatial aspects is a precondition for assessing living standards, levels of social security and social equity in dividing material values produced by the society. Statistical analyses of the distribution of wages of the population also form the basis for decisions in the area of state budgets and social policies. Direct connection of wages with purchasing power of the population also puts monitoring the level, structure and development of the wage distribution to the forefront in identifying sales opportunities for products of long-term and short-term consumption. The main aim of this paper is the creation of the clusters of 14 regions in the Czech Republic according to the wage level.

We can also use them at different considerations of the level of tax burden, etc. Knowledge of wage distribution can be also used indirectly, for improvement of procedures used in sample surveys of other statistical characteristics that are correlated with wages. There are, for example, household expenditures, their equipment, use of time, buying intentions and the like. The attitudes that respondents hold in various sociological surveys are related to wages as

well. Knowledge of wage distribution allows an assessment of the living standards of the country's population without distinction as a comparison of the living standards of members of various social groups, ethnic minorities or residents of different regions.

The issue of wage inequality between men and women (gender wage gap) is the difference between wage levels of men and women for the same task. Women in the EU earn on average around 16% less than men and this figure varies greatly in individual countries.

2. Review of Literature

For these reasons, the theme of labour market analysis and related wages and incomes of the population and the issue of gender wage gap are constantly current, and they are the subject of research of many domestic and foreign authors. For example, Alfarhan (2015) analyses the difference in remuneration between men and women in Jordan, based on data from the years 2002, 2006 and 2008 concerning the expenditures and incomes of households. Bárány (2016) demonstrates the relationship between the development of the minimum wage and wage inequality. Bartošová and Želinský (2013) briefly describe the history of attempts to measure poverty before the split of Czechoslovakia. They focus further detailed analyses on monetary poverty, relative material deprivation and subjective perception of poverty already existing in these two divided countries on the basis of microdata from the EU-SILC statistical survey. Dahl, Maire and Munch (2013) study how the decentralization of wage bargaining in the direction from the sector to enterprise affects wage level and wage variability. This research is based on the use of detailed panel data covering the period of decentralization in the Danish labour market. Fisher, Johnson and Smeeding (2015) examine the distributions of income and consumption in the United States using one set of data and acquire the characteristics of both, income and consumption from the same set of individuals. Kolev and Robles (2015) analyse the wage disparity by gender in Peru in 2005–2011 using data from a national survey. Liberati (2015) describes the development of income and income inequality in the world from 1970 to 2009 using the Gini coefficient. Malá (2015) deals with the construction of multidimensional probabilistic models of income distribution of Czech households. Schünemann, Lechner and Wunsch (2015) assess a wage subsidy program aimed at long-term unemployed workers in Germany.

3. Database

Data for this research comes from the official website of the Czech Statistical Office (CSO). The database for this research represents the wage distribution in total, covering all employees of the Czech Republic divided into regions for the period 2009–2015. In individual regions, employees were divided into groups of men and women and the gender wage gap was researched. There is annual data related to gross monthly nominal wages in the year. For example, average wage (median wage) then represents average (median) gross monthly wage during the year.

There are also data in the form of interval frequency distribution with unevenly wide intervals and open extreme intervals. More detailed data or individual data are not currently available. Because CSO only provides data regarding the nominal wage, the conversion of obtained average nominal wages and median nominal wages to the average real wages and median real wages was carried out using the rate of inflation. CSO provides data on the development of the inflation rate in the period.

Differences in average and median wage (nominal and real) between men and women were calculated and the development of these differences was examined. The research variables are gross monthly nominal and real wage and the differences in average and median wage

(nominal and real) between men and women. Data on average gross monthly (nominal) wages in selected individual European countries come from the official website of Eurostat.

Because only data on nominal wages were taken from the website of the CSO, it was necessary to use inflation rates to recalculate it to a real wage that reflects the purchasing power of participants to allow a comparison of wage development for the period researched without the impact of inflation. The rate of inflation is based on the consumer price index, which is a Laspeyres price index. The real wage was calculated using the real wage index, which is calculated as nominal wage index divided by the consumer price index (index of living costs). The data was processed using the SAS and Statgraphics statistical program packages and Microsoft Excel spreadsheets.

Data for this research includes employees in the business and non-business spheres in the Czech Republic. Wages include employees for work performed in the private (business) sphere and salaries in the budget (national, public, non-business) sector. In terms of the data presented on the website of CSO, the term wages include both wages in the business sphere and salaries in the non-business sector.

4. Theory and Methods

4.1 Cluster Analysis

Cluster analysis was used to allocate the Czech regions into relatively homogenous groups according to the level of gross monthly wages in these regions. Multivariate statistical data analysis, which is often used for processing economic data, may include other multivariate methods of statistical data analysis, namely canonical correlation analysis. Makhalova and Pecáková (2015) or Řezanková and Löster (2013) deal with the special aspects of cluster analysis.

One possibility of the utilization of the information contained in multidimensional observations is classification of the set of objects into several relatively homogenous clusters. We have a data matrix X of the type n X p, where n is the number of objects and p is the number of variables. We consider various decompositions S(k) of the set of n objects into k clusters. We look for those decompositions considered the most appropriate. The goal consists in finding objects within the clusters that are similar as far as possible. We here concede only decompositions with disjunctive clusters. This method is described in detail in Everitt, Landau, Leese and Stahl (2011).

4.2 Lognormal Curves

Theoretical aspect of lognormal distribution is amply processed in the statistical professional literature in terms of the course of lognormal curve, its characteristics, methods of point parameter estimation and construction of curve, see for example Johnson, Kotz and Balakrishnan (1994) or Kleiber and Kotz (2003). The importance of lognormal distribution as a model for sample distribution cannot be questioned. This model has found an application in various fields from astronomy, through technology, economics up to sociology. The characteristic features of the process captured using lognormal model are sequential effect of interdependent factors, the tendency to development in geometric sequence and overgrowth of random variability in systematic variability, i.e. differentiation. In the field of economics, wages and incomes of the population are among the many phenomena that the lognormal model allows to interpret.

4.3 Other Methods Used

The substance of the maximum likelihood method of point parameter estimation is explained in detail in statistical literature, for example in Johnson, Kotz and Balakrishnan (1995). The issue of L-moments is discussed for example in Hosking (1990) or Kyselý and Picek (2007).

Some statistical model of data is supposed. We involve L as the maximum value of the likelihood function for this model, k represents the number of parameters estimated and n is sample size. Akaike information criterion (AIC) has the form

$$AIC = 2k - 2\ln L,$$ (1)

and Bayesian information criterion (BIC) is defined

$$BIC = k\ln n - 2\ln L.$$ (2)

From a number possible models, the preferred model is that with minimal AIC or BIC values. AIC and BIC also include a penalty that is an increasing function of the number of estimated parameters.

Simple descriptive measurements were used for characterizing wage distribution, see for example Barber (1988) or Triola (1989).

In terms of development of gender wage inequality and for its forecast, the trend analysis of time series was used. Interpolation criteria were used for choice the appropriate trend function. The substance of the analysis of economic time series used is explained in Brockwell and Davis (2002) or Cowperwait and Metcalfe (2009).

5. Results

Figures 1 and 2 provide an information of the placement and level of gross monthly wage according to the individual regions in the Czech Republic (Table 1 presents an official designation of the individual regions in the Czech Republic). We can see from these figures substantial higher wage level in Capital Prague Region. We can still observe a relatively high level of wages in Central Bohemian Region and Pilsen Region. A relatively low level of wages is seen in Karlovy Vary Region, Zlin Region and Olomouc Region.

Capital Prague Region is not self-sufficient by food or energy. However, in macroeconomic statistics it considerably overtakes other domestic regions. Prague wages seem as if they belonged to another, wealthier country. The reason may be the higher productivity of local residents, focused on progressive branches of finance and informatics. Politics of the firms may also perform some role, because they often produce values in regions outside the Capital Prague Region, but the profits are added up in the capital, where they have their headquarters. As expected, the three regions with the highest wage levels are simultaneously the three counties with the lowest level of unemployment rate in the same order. However, the order of regions according to the lowest wage level is not the same as their order by the highest unemployment rate. There are the Moravian-Silesian Region, the Usti Region, neither of them belonging between the triplet of regions with the lowest wage level, and the Karlovy Vary Region; see in detail Table 2.

Figures 3–6 provide an overview of the results of cluster analysis of the Czech regions (into three or five clusters) according to the wage level using method of the furthest neighbour and Euclidean distance metric.

Whether it is the average or median monthly wage (three or five clusters), the first cluster always contains only one element, namely the Capital Prague Region, which is due to markedly higher wage levels in this region than in other regions of the Czech Republic.

Figure 1: Placement of individual regions and average gross monthly wage in the Czech Republic in 2015 by regions
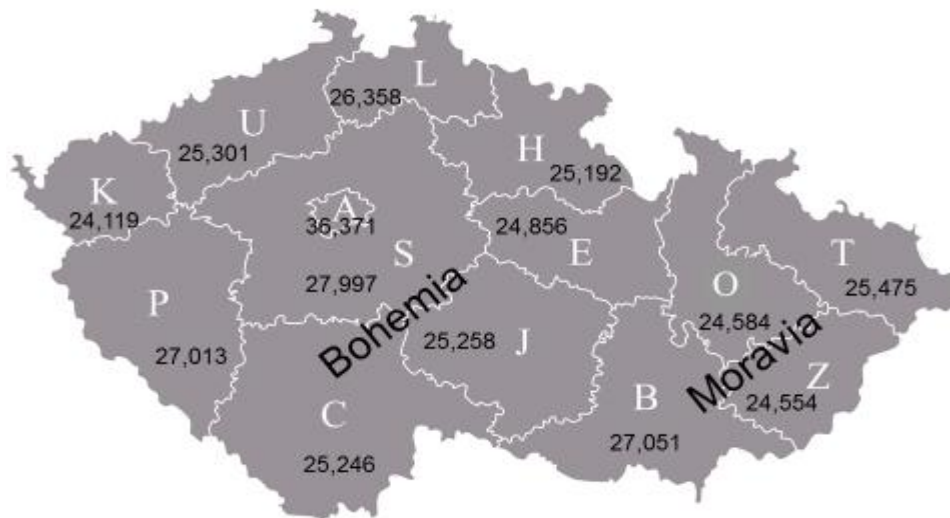
Source: www.czso.cz; own adjustment

Figure 2: Placement of individual regions and median gross monthly wage in the Czech Republic in 2015 by regions
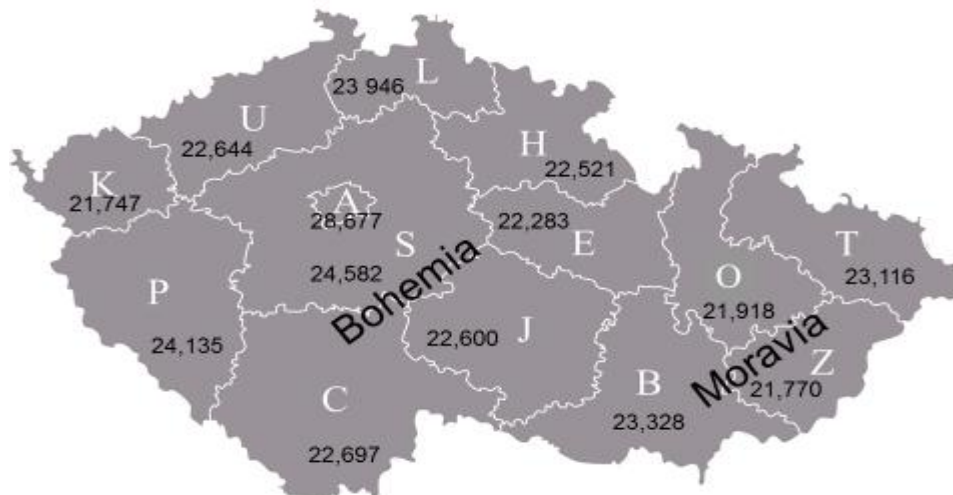
Source: www.czso.cz; own adjustment

In terms of the separation of regions of the Czech Republic into only three clusters by average wage, the second cluster has four-elements. There are Central Bohemian Region, Pilsen Region, Liberec Region and South-Moravian Region. However, according to the median wage, the second cluster has five elements. Apart from the four counties, this cluster also paradoxically contains Moravian-Silesian region, where general unemployment rate

achieves the highest value from the whole Czech Republic. The third cluster then contains the remaining regions.

Table 1: Official designation[1]) of the individual regions in the Czech Republic[2])

| Region	Indication	Region	Indication
Capital Prague Region	A	Hradec Kralove Region	H
Central Bohemian Region	S	Pardubice Region	E
South Bohemian Region	C	Vysocina Region	J
Pilsen Region	P	South Moravian Region	B
Karlovy Vary Region	K	Olomouc Region	O
Usti Region	U	Zlin Region	Z
Liberec Region	L	Moravian-Silesian Region	T

Source: www.mdcr.cz

Table 2: Annual average of general unemployment rate (in %) in individual regions of the Czech Republic in 2015

	Region													
	Central Prague Region	Central Bohemian Region	South Bohemian Region	Pilsen Region	Karlovy Vary Region	Usti Region	Liberec Region	Hradec Kralove Region	Pardubice Region	Vysocina Region	South Moravian Region	Olomouc Region	Zlin Region	Moravian-Silesian Region
Unempl. rate	2.8	3.5	4.0	3.8	6.7	7.6	5.5	5.6	4.6	4.7	5.0	5.9	4.7	8.1

Source: www.czso.cz

Figure 3: Cluster analysis using three clusters, method of the furthest neighbour and Euclidean distance metric, average wage in 2015
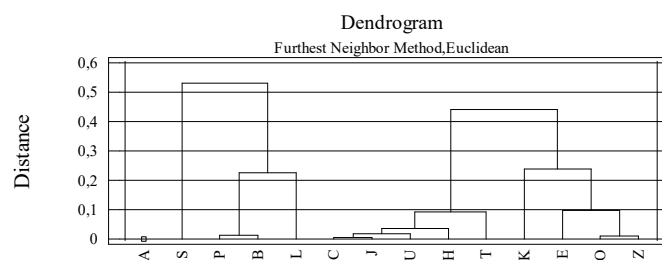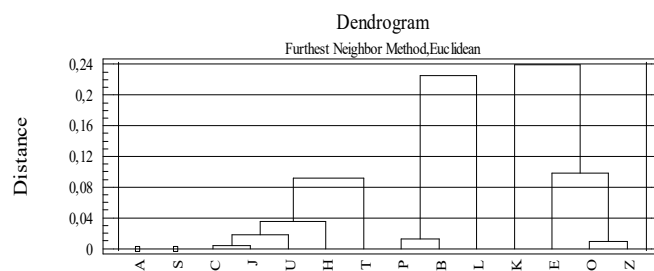

Source: Own research

According to the separation of regions of the Czech Republic into five clusters by average wage, the second cluster contains only one element, namely the Central Bohemian Region. Nevertheless according to the median wage, the second cluster has three elements, specifically the Central Bohemian Region, Pilsen Region and Liberec Region. The third cluster always

---

[1]) Designation of individual regions is derived from the name of the appropriate county town.

[2]) Regions on a dark background belong to Bohemia, regions on a white background belong to Moravia.

contains five elements. These are the South Bohemian Region, Usti Region, Hradec Kralove Region, Vysocina Region and Moravian-Silesian Region by average wage and South Bohemian Region, Usti region, Hradec Kralove Region, Pardubice Region and Vysocina Region, according to the median wage. The fourth cluster has only three elements in both cases (average and median wage). According to the average wage, these are the Pilsen region, Liberec region and South Moravian Region and there are Karlovy Vary Region, Olomouc Region and Zlin Region by median wage (regions with the lowest wage levels). In the case of average wage, the fifth cluster contains the remaining four regions, and in the case of median wage, it contains only two remaining regions (see in detail Figures 3–6).

Figure 4: Cluster analysis using five clusters, method of the furthest neighbour and Euclidean distance metric, average wage in 2015
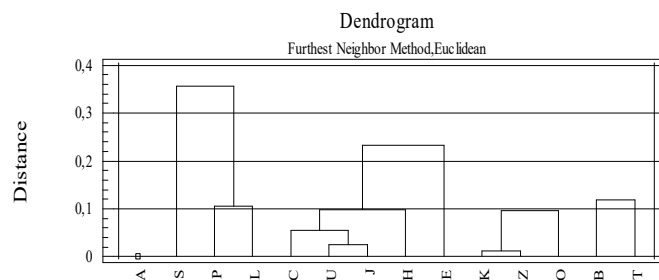

Source: Own research

Figure 5: Cluster analysis using three clusters, method of the furthest neighbour and Euclidean distance metric, median wage in 2015
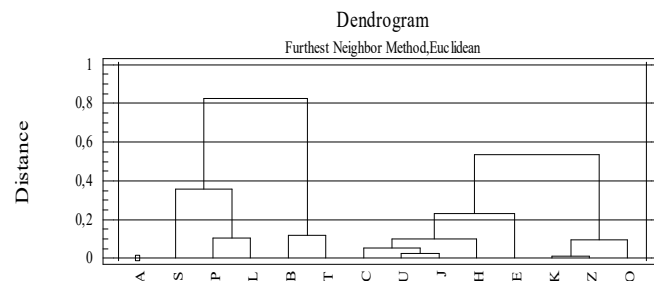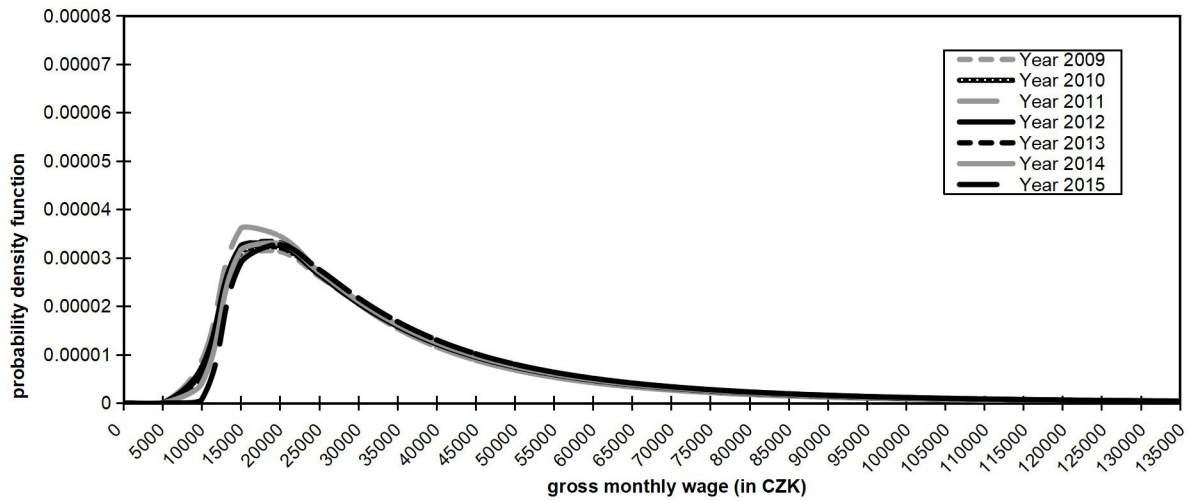

Source: Own research

Figure 6: Cluster analysis using five clusters, method of the furthest neighbour and Euclidean distance metric, median wage in 2015
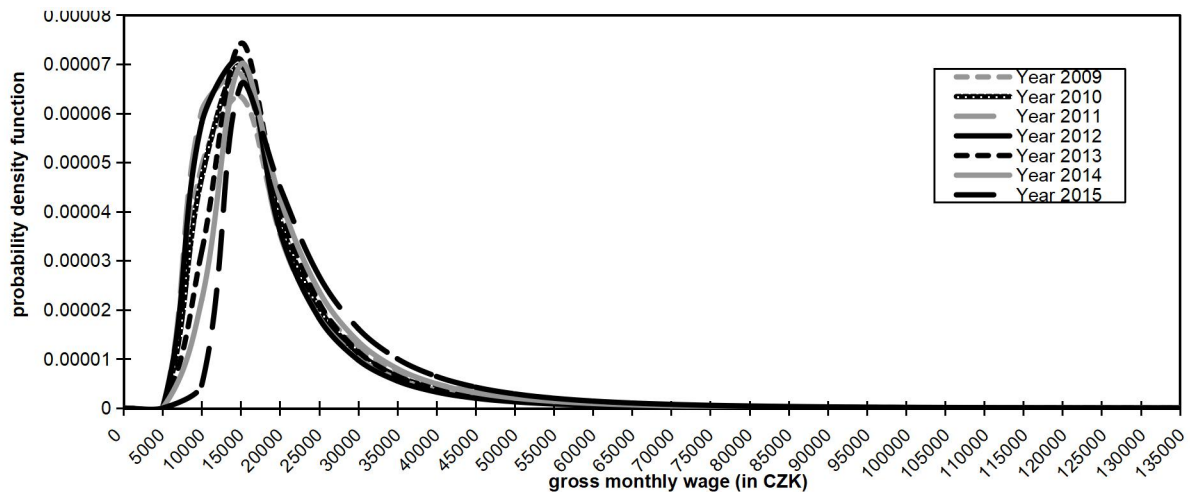

Source: Own research

Figure 7: Development of model wage distributions − Capital Prague Region

Source: Own research

Figure 8: Development of model wage distributions − Karlovy Vary Region

Source: Own research

Table 3: Development of the minimum wage (in CZK) since 2009

Year	2009	2010	2011	2012	2013[3]	2014	2015	2016	2017
Minimum wage	8,000	8,000	8,000	8,000	8,000[4] 8,500[5]	8,500	9,200	9,900	11,000

Source: www.mpsv.cz

Since 2009 the theoretical wage model has been constructed for wage distribution of each region. These models of wage distribution are based on use of the probability density function of the three-parametric lognormal curves, whose parameters were estimated using the maximum likelihood method. The beginning of these curves is represented by the value of

---

[3] In 2013, the beginning of lognormal curves was stated in proportion, i.e.

$$8\ 000 + 7 * \frac{8\ 500 - 8\ 000}{12} \doteq 8\ 292 .$$

[4] Since 1st January 2013 to 31st July 2013.

[5] Since 1st August 2013 to 31st December 2013.

minimal wage in the years, see Table 3. The accuracy of these models obtained was compared using Akaike information criterion and Bayesian information criterion, both of which take number of parameters of the corresponding wage model into consideration.

Table 4: Estimations of parameters of three-parametric lognormal distribution using the maximum likelihood method (parameter θ is the minimum wage in the year)

Region	Est.	Year						
		2009	2010	2011	2012	2013	2014	2015
Capital Prague Region	μ	9,937	9,911	9,803	9,885	9,875	9,885	9,907
	σ^2	0,765	0,746	0,740	0,761	0,752	0,765	0,777
Central Bohemian Region	μ	9,426	9,443	9,342	9,366	9,412	9,467	9,512
	σ^2	0,732	0,707	0,683	0,702	0,699	0,711	0,742
South Bohemian Region	μ	9,112	9,156	9,124	9,092	9,176	9,250	9,312
	σ^2	0,701	0,688	0,633	0,633	0,627	0,632	0,668
Pilsen Region	μ	9,298	9,316	9,202	9,226	9,275	9,371	9,426
	σ^2	0,653	0,634	0,639	0,649	0,630	0,639	0,682
Karlovy Vary Region	μ	9,113	9,054	8,978	8,963	9,060	9,149	9,239
	σ^2	0,737	0,627	0,686	0,635	0,577	0,616	0,670
Usti Region	μ	9,264	9,280	9,130	9,169	9,216	9,266	9,337
	σ^2	0,711	0,674	0,671	0,664	0,637	0,654	0,707
Liberec Region	μ	9,326	9,292	9,146	9,182	9,247	9,308	9,392
	σ^2	0,950	0,656	0,636	0,625	0,618	0,626	0,689
Hradec Kralove Region	μ	9,144	9,201	9,087	9,139	9,190	9,259	9,304
	σ^2	0,640	0,671	0,623	0,625	0,620	0,636	0,661
Pardubice Region	μ	9,225	9,168	9,105	9,118	9,160	9,231	9,306
	σ^2	0,920	0,697	0,642	0,660	0,642	0,650	0,704
Vysocina Region	μ	9,195	9,204	9,093	9,133	9,193	9,255	9,322
	σ^2	0,754	0,691	0,616	0,640	0,605	0,627	0,685
South Moravian Region	μ	9,358	9,394	9,268	9,311	9,361	9,404	9,460
	σ^2	1,028	1,025	1,028	1,027	1,018	1,014	1,009
Olomouc Region	μ	9,204	9,203	9,088	9,086	9,164	9,239	9,296
	σ^2	0,661	0,659	0,632	0,657	0,651	0,633	0,719
Zlin Region	μ	9,077	9,139	9,075	9,065	9,149	9,215	9,277
	σ^2	0,739	0,704	0,656	0,680	0,626	0,633	0,686
Moravian-Silesian Region	μ	9,196	9,251	9,220	9,233	9,262	9,290	9,345
	σ^2	0,681	0,664	0,665	0,662	0,656	0,654	0,702

Source: Own research

Table 5: The values of Akaike and Bayesian information criteria

Region	Crit.	Year						
		2009	2010	2011	2012	2013	2014	2015
Capital Prague Region	AIC	1,715,847	1,605,510	1,514,319	1,456,900	1,362,167	1,291,967	1,219,753
	BIC	1,715,870	1,605,533	1,514,342	1,456,923	1,362,191	1,291,990	1,219,776
Central Bohemian Region	AIC	544 ,441	536,317	528,511	543,060	546,067	556,645	577,484
	BIC	544,462	536,338	528,533	543,082	546,089	556,666	577,505
South Bohemian Region	AIC	297,991	296,670	282,364	284,622	285,256	289,311	303,276
	BIC	298,011	296,691	282,384	284,643	285,277	289,331	303,296
Pilsen Region	AIC	269,898	267,230	272,194	278,447	275,599	281,651	298,363
	BIC	269,919	267,250	272,215	278,468	275,619	281,671	298,383
Karlovy Vary Region	AIC	127,711	115,055	123,387	117,626	110,162	116,655	124,583
	BIC	127,730	115,074	123,406	117,644	110,180	116,674	124,602
Usti Region	AIC	348,582	336,708	336,174	334,146	324,956	331,734	350,898
	BIC	348,603	336,728	336,195	334,167	324,977	331,755	350,919
Liberec Region	AIC	231,143	185,448	183,454	183,282	183,845	187,689	203,032
	BIC	231,162	185,468	183,474	183,302	183,864	187,709	203,052
Hradec Kralove Region	AIC	240,158	250,588	239,702	242,212	242,731	249,249	258,341
	BIC	240,178	250,609	239,722	242,233	242,751	249,270	258,361
Pardubice Region	AIC	275,403	235,724	226,401	234,568	233,737	239,552	257,057
	BIC	275,423	235,744	226,421	234,588	233,757	239,572	257,077
Vysocina Region	AIC	238,726	228,425	213,311	222,088	215,872	224,336	241,838
	BIC	238,746	228,445	213,331	222,108	215,892	224,356	241,858
South Moravian Region	AIC	786,921	788,918	793,299	796,116	795,440	796,875	797,641
	BIC	786,943	788,940	793,321	796,138	795,462	796,897	797,663
Olomouc Region	AIC	255,008	261,376	260,425	274,584	279,522	280,684	313,825
	BIC	255,028	261,396	260,445	274,604	279,543	280,705	313,846
Zlin Region	AIC	289,572	282,403	271,177	279,996	265,977	270,105	288,029
	BIC	289,593	282,423	271,197	280,017	265,997	270,125	288,050
Moravian-Silesian Region	AIC	581,796	574,637	578,439	579,575	578,706	580,905	613,021
	BIC	581,817	574,659	578,461	579,597	578,728	580,927	613,042

Source: Own research

Theoretical wage models using three-parametric lognormal curves and maximum likelihood method of parametric estimation were constructed. Table 4 shows values of parameters estimated and Table 5 presents the values of Akaike and Bayesian information criteria, which enable to assess the accuracy of the estimation. Theoretical wage models of region with the highest wage level (Capital Prague Region) and region with the lowest wage level (Karlovy Vary Region), are presented in Figures 7 and 8.

These figures enable some comparison of the development of wage distribution of the regions with the highest wages on the one hand and with the lowest wages on other hand during the last seven years. We can see from these figures that wage distributions with higher wage level are also distinguished by higher variability than those with lower wage level. In

addition, wage distributions with lower level of wage are more skewed and they have also higher kurtosis than those with higher wage level, see Figures 7 and 8.

We can see from Table 5 that wage models for the Capital Prague Region show the lowest accuracy. On the other hand, we can observe from this table that Karlovy Vary Region with the lowest wage level represents simultaneously the highest accuracy of wage models. It is necessary to point out here that the both information criteria (AIC and BIC) take in their account the number of parameters of the model. It is apparent from Table 5 that the model accuracy is related to the wage level in the region. We can pronounce in the principle that the model accuracy is the worst, the wage level in the region is higher and vice versa.

Trend model of gender wage gap was carried out using the parametric manner when appropriate trend function was chosen using interpolation criteria (quadratic trend function in all cases), and the nonparametric manner using moving averages, which represent one of the methods of adaptive approach to modelling the trend component. The moving averages of three lengths were used regard to the length of the time series researched. Figure 9 represents the same development of the gender wage gap and its models for Central Bohemian Region.

The suitability of a selected trend function was subsequently verified using the individual t-test and overall F-test. All of the tests appeared statistically significant at the 5% significance level. The significant autocorrelation is not in the residues of the models, as well as all values of Durbin-Watson statistics was found in the interval (1.4, 2.6), so they are equal to two values and thus they speak in advantage of the independence of random failures. Also calculated values of the coefficient of determination spoke to the benefit of quadratic trend functions gained.

The extrapolation of the time series studied in the sense that we shorten the time series examined by a number of observations, which will be then predict using the selected model and we compare the accuracy of the extrapolation for example using Theil coefficient of mismatch, was not performed due to the short time series. In the case of gender wage gap in real average and median wage, it proceeded in exactly the same way.

Figure 9: Difference between the levels of male and female wages: Central Bohemian Region
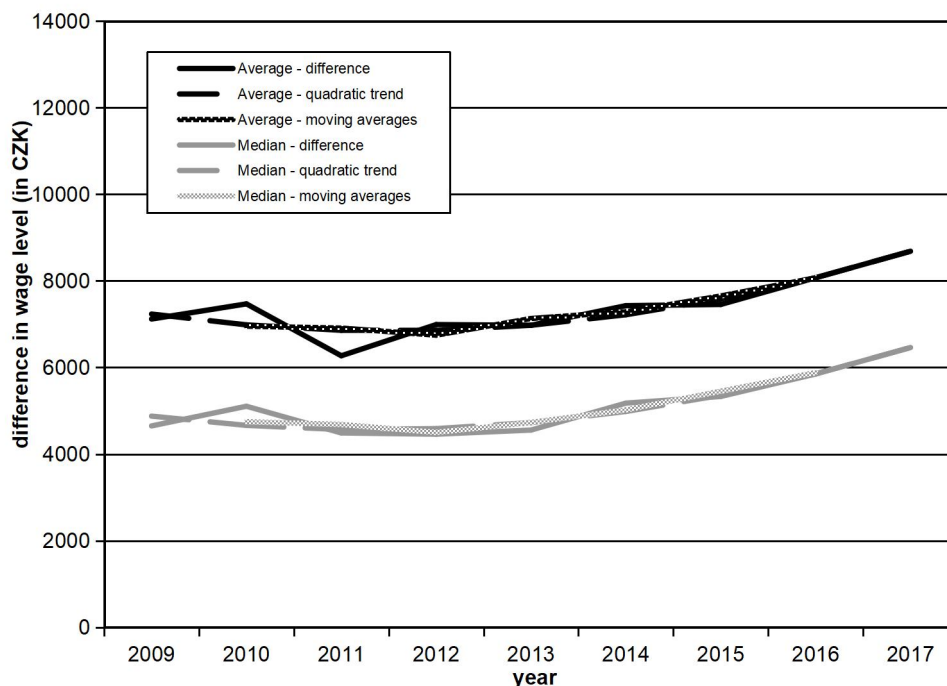
Source: Own research

In the vast majority of the regions, the gender wage gap had a decreasing tendency at the beginning of the researched period, probably as a result of the global economic crisis. However, after its unwinding, the gender wage gap grows again. Only in Usti Region and partly in South Bohemian Region and Moravian-Silesian Region, this is not true and the gender wage gap fails to grow during recent years in these regions. Similar conclusions were drawn on the basis of real wages.

Based on these quadratic trend functions, the predictions of gender wage gap in average and median (nominal) wage were created by individual regions for 2016 and 2017, see Table 6.

Table 6: Predictions of differences of wage level of men and women (in CZK) in the Czech Regions for 2016 and 2017

| Region	Average		Median	
	Year 2016	Year 2017	Year 2016	Year 2017
Capital Prague Region	11,607	13,074	5,866	6,931
Central Bohemian Region	8,071	8,683	5,850	6,460
South Bohemian Region	5,649	6,162	3,623	3,684
Pilsen Region	6,765	7,650	4,535	5,014
Karlovy Vary Region	5,812	6,774	4,913	5,627
Usti Region	5,329	5,415	3,549	3,328
Liberec Region	7,851	9,652	4,795	5,546
Hradec Kralove Region	5,312	5,726	3,580	3,913
Pardubice Region	6,384	7,643	4,505	5,188
Vysocina Region	7,796	8,672	6,019	6,631
South Moravian Region	7,572	8,443	4,630	5,204
Olomouc Region	6,180	7,256	4,542	5,131
Zlin Region	7,172	7,740	5,372	5,794
Moravian-Silesian Region	5,986	6,058	5,232	5,503

Source: Own research

6. Discussion

Figures 10–12 illustrate the above statements using the results of cluster analysis for 31 selected European countries. These states were divided into three, five and seven clusters according to their average gross monthly (nominal) wage in 2015 using method of the furthest neighbour and Euclidean distance metric.

Figure 10 shows the segmentation of these countries into three clusters. The first cluster consists of countries with medium wage levels in Europe. There are the following fourteen states: Austria, Belgium, Finland, France, Germany, Iceland, Ireland, Italy, Luxembourg, Netherlands, Spain, Sweden, Switzerland and United Kingdom. The second cluster consists of states with rather lower wage levels, to where the countries of the former socialist block (Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia) belong, inter alia (Cyprus, Greece, Malta and Portugal). Two countries with the highest wage levels make up the third cluster. There are Denmark and Norway.

Figure 11 represents the classification of these states into five clusters. The first cluster is formatted by states with rather higher levels of nominal wage, but not with the countries with the highest wage levels in Europe. These are Austria, Belgium, Finland, France, Germany Ireland, Luxembourg, Netherlands, Sweden and Switzerland. It is interesting that the whole of the second cluster is represented by the countries of the former socialist block with the

exception of Slovenia: Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania and Slovakia. The third cluster is made up of rather poorer countries of the original "Western" Europe (Cyprus, Greece, Malta and Portugal) plus Slovenia. Two states with the highest levels of nominal wages, Norway and Denmark, represent the fourth cluster. Four countries; Iceland, Italy, Spain and United Kingdom, represent the fifth cluster.

Figure 10: Cluster analysis using three clusters, method of the furthest neighbour and Euclidean distance metric, average wage in 2015
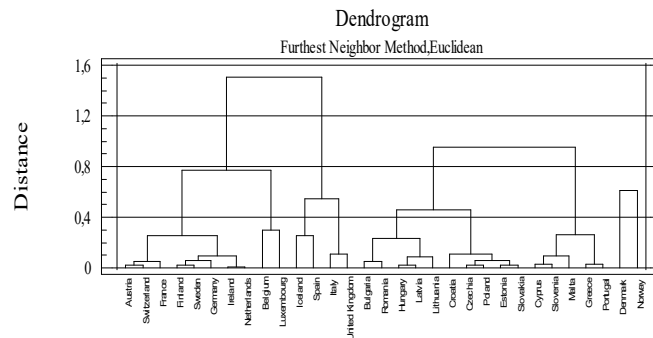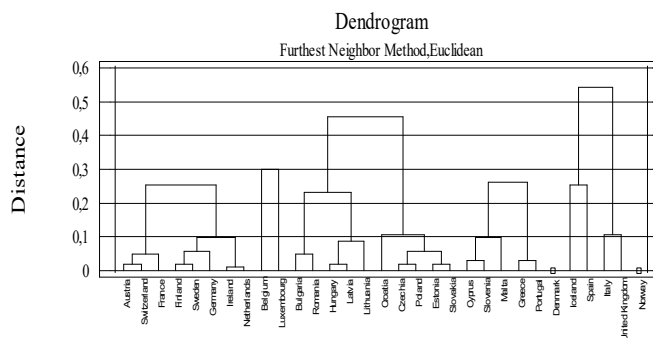

Source: Own research

Figure 11: Cluster analysis using five clusters, method of the furthest neighbour and Euclidean distance metric, average wage in 2015
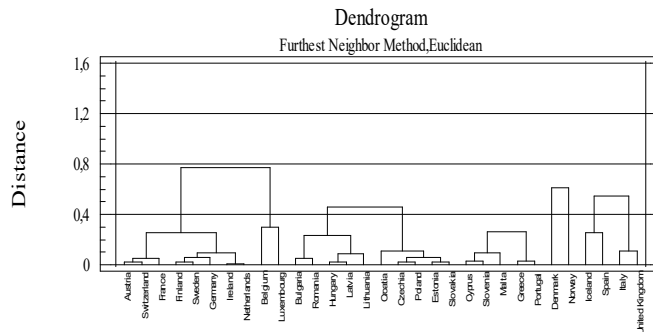
Source: Own research

Figure 12: Cluster analysis using seven clusters, method of the furthest neighbour and Euclidean distance metric, average wage in 2015


Source: Own research

Figure 12 represents the allocation of European countries into seven clusters. Eight advanced countries of Western and Northern Europe constitute the first cluster. There are Austria, Finland, France, Germany, Ireland, Netherlands, Sweden and Switzerland. Only two states with very high level of nominal wages; Belgium and Luxembourg, form the second cluster. European countries of the former socialist block, with the exception of Slovenia, represent the entire third cluster. There are Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania and Slovakia. Cluster number four is constituted by somewhat less developed countries of the Western and Southern Europe, such as Cyprus, Greece, Malta, Portugal and Slovenia. The fifth and the seventh clusters have only one country with the highest level of nominal wages. Denmark constitutes the fifth cluster and Norway the seventh cluster. The remaining four countries, such as Iceland, Italy, Spain and United Kingdom form the sixth cluster.

7. Conclusion

People in the Central Prague Region have the highest wages and people in the Karlovy Vary Region have, on the contrary, the lowest wages. The average gross monthly wage of the employees in the Central Prague Region was 36,371 CZK in 2015, and in the Karlovy Vary Region it was only 24,119 CZK in the same year. Dwellers of the Central Bohemian region have relatively high wages, which average was 27,997 CZK in 2015. Pilsen and South Moravian Region are characterized by relatively high wage levels, too. There are on average 27,013 CZK in Pilsen Region and 27,051 in South Moravian Region in 2015. However, counties with high wages are also at the same time characterized by relatively high gender wage gaps.

Figure 13: Europe's purchasing power index per inhabitant in 2015
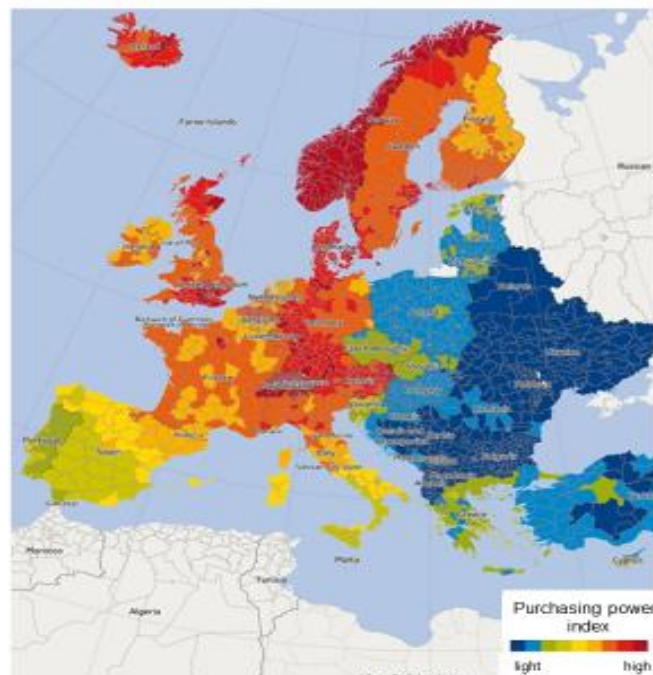
Source: www.gfk-geomarketing.com, own customization

Purchasing power of the Czech population currently reaches less than 60% of the European average. Purchasing power of the Central and Eastern Europeans annually rose by 5%.

However, division into Western European countries on the one hand and Eastern European countries on the other hand is still holding.

The Czech Republic also economically lags behind Western European countries and the idea that it will catch up with them in the near future falls into the category of wishful thinking. Figure 13 offers a view of the old continent using the optics of economic strength of individual European regions, where the border between these two blocks mentioned is clearly visible. The differences within Europe are evident: Western European countries against Eastern European countries, but also Northern European countries against Southern European countries, see Figure 13.

Acknowledgements

References

[1] Alfarhan, U. F. 2015. Gender Earnings Discrimination in Jordan: Good Intentions Are not Enough. International Labour Review, vol. 154, iss. 4, pp. 563–580.

[2] Bárány, Z. L. 2016. The Minimum Wage and Inequality: The Effects of Education and Technology. Journal of Labour Economics, vol. 34, iss. 1, pp. 237–274.

[3] Barber, G. M. 1988. Elementary Statistics for Geographers. New York: The Guilford Press. ISBN 0-89862-777-X.

[4] Bartošová, J., Želinský, T. 2013. The Extent of Poverty in the Czech and Slovak Republics 15 Years after the Split. Post-Communist Economies, vol. 25, iss. 1, pp. 119–131.

[5] Brockwell, P. J., Davis, R. A. 2002. Introduction to Time Series and Forecasting. New York: Springer. ISBN 978-0-387-95351-9.

[6] Cowperwait, P. S. B., Metcalfe, A. V. 2009. Introductory Time Series with R. New York: Springer. ISBN 978-0-387-88697-8.

[7] Dahl, C. M., Maire, D., Munch, J. R. 2013. Wage Dispersion and Decentralization of Wage Bargaining. Journal of Labour Economics, vol. 31, iss. 3, pp. 501–533.

[8] Everitt, B. S., Landau, S., Leese, M., Stahl, D. 2011. Cluster analysis. London: John Wiley & Sons. ISBN 978-0-470-74991-3.

[9] Fisher, J., Johnson, D. S., Smeeding, T. M. 2015. Inequality of Income and Consumption in the U.S.: Measuring the Trends in Inequality from 1984 to 2011 for the Same Individuals. Review of Income and Wealth, vol. 61, iss. 4, pp. 630–650.

[10] Hosking, J. R. M. 1990. L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of the Royal Statistical Society (Series B), vol. 52, iss. 1, pp. 105–124.

[11] Johnson, N. L., Kotz, S., Balakrishnan, N. 1994. Continuous Univariate Distributions 1. New York: John Wiley & Sons. ISBN 0-471-58495-9.

[12] Johnson, N. L., Kotz, S., Balakrishnan, N. 1995. Continuous Univariate Distributions 2. New York: John Wiley & Sons. ISBN 0-471-58494-0.

[13] Kleiber, C., Kotz, S. 2003. Statistical Size Distributions in Economics and Actuarial Sciences. New York: John Wiley & Sons. ISBN 0-471-15064-9.

[14] Kolev, A., Robles, P. S. 2015. Ethnic Wage in Peru: What Drives the Particular Disadvantage of Indigenous Women? International Labour Review, vol. 154, iss. 4, pp. 417–448.

[15] Kyselý, J., Picek, J. 2007. Regional Growth Curves and Improved Design value Estimates of Extreme Precipitation Events in the Czech Republic. Climate research, vol. 33, iss. 3, pp. 243−255.

[16] Liberati, P. 2015. The World Distribution of Income and Its Inequality, 1970–2009. Review of Income and Wealth, vol. 61, iss. 2, pp. 248–273.

[17] Makhalova, E., Pecáková, I. 2015. The Fuzzy Clustering Problems and Possible Solutions. In: The 9th International Days of Statistics and Economics – MSED. Prague: University of Economics, Prague. ISBN 978-80-87990-06-3.

[18] Malá, I. 2015. Multivariate Probability Model for Incomes of the Czech Households. Politická ekonomie, vol. 63, iss. 7, pp. 895–908.

[19] Řezanková, H., Löster, T. 2013. Shluková analýza domácností charakterizovaných kategoriálními ukazateli (Cluster Analysis of Households Characterized by Categorical Indicators). E+M Ekonomie a Management, vol. 16, iss. 3, pp. 139–147.

[20] Schünemann, B., Lechner, M., Wunsch, C. 2015. Do Long-Term Unemployed Workers Benefit from Targeted Wage Subsidies? German Economic Review, vol. 16, iss. 1, pp. 43–64.

[21] Triola, M. F. 1989. Elementary Statistics. Redwood City: The Benjamin/Cummings Publishing Company. ISBN 0-8053-0271-9.