

Wage Level Comparison in the OECD Member States

Diana BÍLKOVÁ

University of Economics, Prague, Czech Republic
bilkova@vse.cz

Abstract. This paper focuses on the analysis of wage levels and other related indicators, such as the minimum wage, GDP per capita and unemployment rate, of 32 selected OECD member countries. The countries were chosen from both the 20 founding states in 1960 and the latter acceding countries, including the former Socialist bloc countries. The main aim of this paper is to create clusters of selected OECD member states that are similar as possible in terms of these variables. Cluster analysis was used for this purpose. No less important objective of this research was to find out, which of these variables affect the wage levels in these countries, including the type of dependency. Special attention has been paid to comparing wage developments in recent years between G7 and V4. We have found that dividing line between Western and Eastern European countries still persists and will likely to remain so for some time to come.

Keywords: Wages and GDP in the OECD Member States, Cluster Analysis, Ward's Method and Euclidean Distance, Stepwise Regression and Forward Selection, Linear Regression Hyperplane, Polynomial Regression.

1 Introduction

All OECD member countries are economically mature. Despite this fact, large differences in citizens' living standards are among them, as evidenced inter alia by the average gross wage. The average wage is overestimated by the wages of the best-paid professionals in all OECD member countries, so the average wage does not correspond to the vision of so called common wage. Wages of the worst-paid employees mostly stagnate. About one third of employees work for average or higher wage, a specific number is different in each country. Scandinavian countries, the Czech and Slovak Republics are among the countries with the lowest wage differences. More employees then achieve average and higher wage in these countries than in especially non-European OECD member states. For this reason, both European and non-European OECD countries have been included into the research. Increase in minimum wage also contributes to reducing wage differentials, when wage restriction comes from below. For this reason, one of explanatory variable was chosen the minimum wage. The highest wage differences are in Mexico and Chile. The highest average wages are in the most economically advanced countries in the world. The wage growth is based on a high degree of personal and economic freedom,

a sophisticated educational system preparing skilled employees, excellent business conditions and functional public administration. If the firms are successful, there is also high labour market supply and capable employees can find more interesting or better-valued jobs.

For this reason, the wage and related indicators issue is still topical and it is matter of interest for many researchers. For example, [1] describes the history of attempts to measure poverty prior to the split. Their analyses are focused on monetary poverty, relative material deprivation and subjective perception of poverty in two countries fifteen years after the split; [3] analyses the development of wages in the Czech Republic by education level; [6] analyses the equivalized total net annual incomes of the Czech households (in CZK) in 2007–2010.

This paper deals with the situation regarding wage levels in 32 chosen OECD member states, which were selected from a total 35 OECD member states. Countries like Iceland, Latvia and Turkey have not been included in the research because of insufficient data in purchasing power parity in USD with constant prices in 2015. Special attention is paid to the context of employee' average annual gross wage with other economic indicators, such as real minimum wage, structural unemployment rate, GDP per head of population. The minimum wage is not set by legislation in some countries, such as Austria, Denmark, Finland, Italy, Norway, Sweden and Switzerland. The minimum wage is considered zero in these countries. In Nordic countries, it is usually negotiated in collective agreements. This research has several aims. Selected OECD member as the objects were clustered into groups of similar countries in term of the above variables. Cluster analysis and within that Ward's method and Euclidean distance were used for construction five, seven, nine and eleven clusters. There are various methods for determining the optimal number of clusters in cluster analysis. However, there is no definitive answer to the question of determining the optimal number of clusters. The problem lies in that cluster analysis is basically an exploratory approach. Linear regression hyperplane was used to research the dependence of the average annual gross wage on the remaining three variables. Normality of the variables was verified using both Kolmogorov-Smirnov goodness of fit test and the visual one. The issue of heteroscedasticity was verified using Glejser test and visual manner (random course of residues). The variables were put into model using stepwise regression and forward selection. Only one variable is suitable in this sense, namely GDP per head of person. Multicollinearity exploration was unnecessary, since only one independent variable was inserted into a model. Polynomial regression of the second stage was found to be better than linear regression. The suitability of the model chosen of dependence of average annual gross wage on GDP per head of person was subsequently verified using both by individual t-tests, through by the overall F-test and using the adjusted determination index. All these results are for 2015.

The main research hypothesis consists in the statement that division into Western European countries on the one hand and Eastern European countries on the other hand is still holding.

2 Database

Table 1 gives an overview of the choice of 32 member states and their two-letter abbreviations according to the norm number ISO 3166-1 alpha-2.

Table 1. Alphabetically arranged two-letter country abbreviations by norm number ISO 3166-1 alpha-2.

Marking	Country	Marking	Country
AU	Australia	IE	Ireland
AT	Austria	IL	Israel
BE	Belgium	IT	Italy
CA	Canada	JP	Japan
CH	Switzerland	KR	South Korea
CL	Chile	LU	Luxembourg
CZ	Czech Republic	MX	Mexico
DE	Germany	NL	Netherlands
DK	Denmark	NO	Norway
EE	Estonia	NZ	New Zealand
ES	Spain	PL	Poland
FI	Finland	PT	Portugal
FR	France	SE	Sweden
GB	Great Britain	SI	Slovenia
GR	Greece	SK	Slovak Republic
HU	Hungary	US	United States

Data for this paper come from the official website of the Organization of Economic Cooperation and Development (OECD), see [9]. There are data of average annual (gross) wage after conversion into purchasing power parity (PPP) in USD with constant prices in 2015, real annual minimum wage after conversion into PPP in USD with constant prices in 2015, structural unemployment rate in percentages and GDP per head of population after conversion into purchasing power parity (PPP) in USD with constant prices in 2015. The conversion to purchasing power parity allows comparing the purchasing power of the population of different countries. Data are for the period since 2000 to 2015 and average annual wage represents the main variable.

The research data include wages and salaries paid to employees for work performed in the private (business) and public (state budget, non-business) sectors, respectively. In terms of the data presented on the CSO website, “wages” cover remuneration for work done in both the sectors. Data were processed using statistical program packets SAS and Statgraphics and Microsoft Excel Spreadsheet.

3 Theory and Methods

3.1 Cluster Analysis

Cluster analysis was used to divide the selected OECD member states into relatively homogeneous groups according to their respective gross monthly wage levels. Multidimensional observations can be used when classifying a set of objects into several relatively homogeneous clusters. We have a data matrix X of $n \times p$ type, where n is the number of objects and p is the number of variables. Assuming various decompositions $S^{(k)}$ of the set of n objects into k clusters, we look for the most appropriate decompositions. The aim is to find the objects within certain clusters that are as similar as possible to those from other clusters. Only decompositions with disjunctive clusters and tasks with a specified number of classes are conceded.

Criteria for Assessing the Quality of Decomposition. The general task is to assess to what extent the cluster analysis aim has been achieved in a given situation, while applying a specific algorithm. Several criteria – decomposition functions – are proposed for this purpose. The most frequently used ones exhibit the following characteristics. They are the matrices of internal cluster variance

$$\mathbf{E} = \sum_{h=1}^k \sum_{i=1}^{n_h} (\mathbf{x}_{hi} - \bar{\mathbf{x}}_h) \cdot (\mathbf{x}_{hi} - \bar{\mathbf{x}}_h)' \quad (1)$$

and between-cluster variance

$$\mathbf{B} = \sum_{h=1}^k n_h \cdot (\bar{\mathbf{x}}_h - \bar{\mathbf{x}}) \cdot (\bar{\mathbf{x}}_h - \bar{\mathbf{x}})' \quad (2)$$

whose sum is the matrix of total variation

$$\mathbf{T} = \sum_{h=1}^k \sum_{i=1}^{n_h} (\mathbf{x}_{hi} - \bar{\mathbf{x}}) \cdot (\mathbf{x}_{hi} - \bar{\mathbf{x}})' \quad (3)$$

There are vectors of the observations for the i^{th} object and h^{th} cluster \mathbf{x}_{hi} , the averages for the h^{th} cluster $\bar{\mathbf{x}}_h$ and those for the total set $\bar{\mathbf{x}}$. There are p^{th} -membered vectors, \mathbf{E} , \mathbf{B} and \mathbf{T} being symmetric square matrices of the p^{th} order. The principal aim, consisting in the creation of mutually distant compact clusters, is fulfilled by reaching the minimum of the total sum of the deviation squares of all values of corresponding cluster averages

$$C_1 = \text{st } \mathbf{E} = \sum_{h=1}^k \sum_{i=1}^{n_h} \sum_{j=1}^p (x_{hij} - \bar{x}_{hj})^2 \quad (4)$$

i.e. the Ward criterion. Since the $\text{st } \mathbf{T}$ is the same for all decompositions, the minimization of the $\text{st } \mathbf{E}$ means the same as that of the $\text{st } \mathbf{B}$. In order to become independent on the used units of measurement (or, more generally, the invariance to the linear transformations), it is recommended to minimize the determinant of the matrix of the internal cluster variance

$$C_2 = |\mathbf{E}|$$

or to maximize the trace criterion

$$C_3 = \text{st} (BE^{-1}) \text{ or else } C_4 = \text{st} (BT^{-1}).$$

The criteria mentioned above are employed not only retrospectively to assess the decomposition quality accomplished, changes in criterion values also guiding the creation of clusters. Since the criteria ultimately reach the limits (C_1 and C_2 the minimum, C_3 and C_4 the maximum) at $k = n$, it is necessary to find the extreme of the purpose function that properly includes the loss following from the growth in the number of clusters. The Ward criterion, for instance, is proposed to move towards the minimization of the quantity

$$Z_1 = C_1 + z \cdot k, \quad (5)$$

where constant z represents the loss resulting from an increase in the number of clusters by one.

Distance and Similarity of Objects. Having selected the variables characterizing the properties of the clustered objects and found their values, we decided on the method of the evaluation of distance or similarity of objects, the calculation of appropriate measures for all pairs of objects often being the initial stage of clustering algorithm implementation. The symmetric square matrix of $n \times n$ type has zeros or ones on the diagonal, depending on whether it is the matrix of distance D measures or that of similarity A measures, respectively.

Let us now focus on measuring the distance of the objects described by quantitative variables. The Hemming distance can be used when individual variables are roughly on the same level or at least expressed in the same units of measurement

$$D_H(x_i, x_{i'}) = \sum_{j=1}^p |x_{ij} - x_{i'j}|. \quad (6)$$

The Euclidean distance can be applied in the same case

$$D_E(x_i, x_{i'}) = \sqrt{\sum_{j=1}^p (x_{ij} - x_{i'j})^2} \quad (7)$$

as well as the Chebyshev distance

$$D_C(x_i, x_{i'}) = \max_j |x_{ij} - x_{i'j}|. \quad (8)$$

All the above mentioned, measurements have some common drawbacks – the dependence on the used measuring units that sometimes hinders the meaningful acquisition of any sum for different variables and the fact that if the variables are considered in sum with the same weights, the strongly correlated variables have a disproportionately large effect on the outcome. The starting point is the transformation of variables. The adverse effect of the measuring units can be removed by dividing all the values by the balancing factor, which can be presented with the corresponding average \bar{x}_j , standard deviation s_j or the range after deletion of extremes

$$\frac{\max_i x_{ij} - \min_i x_{ij}}{s_j}$$

Particular variables can be also assigned more weight – having decided subjectively or on the basis of relevant information – their values then appearing in the formulas for the calculation of distance

Other measurements of distance and similarity of objects for numerical, ordinal, nominal and alternative variables are described in the professional literature. When dealing with variables of a different type, the Lance-Williams distance is recommended

$$D_{LW}(x_i, x_{i'}) = \frac{\sum_{j=1}^p |x_{ij} - x_{i'j}|}{\sum_{j=1}^p (x_{ij} + x_{i'j})}. \quad (9)$$

Algorithm for the Creation of Hierarchical Sequence of Decompositions. The creation of a hierarchical sequence of decompositions belongs to the most widely used techniques applied in the cluster analysis, occurring sequentially in the following steps:

- **D** matrix calculation of appropriate measurements of distances,
- the start of the decomposition process $S^{(n)}$ from n clusters, each of them containing one object,
- the assessment of the symmetric matrix **D** (a lower or upper triangle), finding two clusters (the h^{th} and h'^{th} ones) whose distance $D_{hh'}$ is minimal,
- the connection of the h^{th} and h'^{th} clusters into a new g^{th} cluster, the replacement of the h^{th} and h'^{th} row and column in the matrix **D** with those of the new cluster, the order of the matrix being reduced by one,
- renumbering of the order of the cycle $l = 1, 2, \dots, n - 1$, the identification of the connected objects h, h' and the level of the connection $d_l = D_{hh'}$,
- returning to step (3) if the creation of decompositions has not been completed by connecting all objects into a single cluster $S^{(1)}$.

A divisive hierarchical procedure, contrary to the agglomerative hierarchical one, is less-used, starting from a single cluster $S^{(1)}$, splitting one of the clusters into two in each step and obtaining $S^{(n)}$ at the end of the process. The results of hierarchical cluster procedures can be effectively displayed in the form of a graphical tree dendrogram.

Given the choice of variables x_1, x_2, \dots, x_p and the matrix of distances **D**, the results of applying the described algorithm vary according to the way the distance between clusters is evaluated.

Nearest Neighbour Method. Within the nearest neighbour method, both clusters, whose connection is considered, are represented by objects that are the closest to each other. The $D_{hh'}$ distance between the h^{th} and h'^{th} clusters therefore represents the minimum of all $q = n_h n_{h'}$ distances between their objects, the procedure of the third phase of the above algorithm thus being specified. In the fourth step, the h^{th} and h'^{th}

rows and columns in the distance matrix are replaced with the new g^{th} cluster's row and column of distances. In the l^{th} cycle, total $n - l - 1$ distances determined by

$$D_{gg'} = \min(D_{g'h}, D_{g'h'}) . \quad (10)$$

can be written.

If the way of evaluation of the proximity or similarity of clusters is given, which also determines the conversion of the distance matrix in each cycle, the above algorithm allows for the creation of a hierarchical sequence of decompositions and construction of the dendrogram.

When using this method, even considerably distant objects can get together in the same cluster if a large number of other objects create a kind of bridge between them. This typical chaining of objects is considered as a drawback, especially if there is a reason for the clusters to acquire the usual elliptical shape with a compact core. This method, however, possesses many positive features that outweigh the above disadvantage.

Farthest Neighbour Method. The method of the farthest neighbour is based on the opposite principle. The criterion for the connection of clusters is the maximum of q possible between-cluster distances of objects. When editing the matrix of distances, we proceed according to

$$D_{gg'} = \max(D_{g'h}, D_{g'h'}) . \quad (11)$$

An adverse chain effect does not occur in this case. On the contrary, there is a tendency towards the formation of compact clusters, not extraordinarily large, though.

Average Linkage Method (Sokal-Sneath Method). As a criterion for the connection of clusters, this method applies an average of the q possible between-cluster distances of objects. When recalculating the distance matrix, we use

$$D_{gg'} = \frac{n_h \cdot D_{g'h} + n_{h'} \cdot D_{g'h'}}{n_h + n_{h'}} . \quad (12)$$

The method often leads to similar results as the farthest neighbour one.

Centroid method (Gower method). Unlike the above methods, this one is not based on summarizing the information on between-cluster distances of objects, the criterion being the Euclidean distance of centroid

$$D_E(\bar{x}_h, \bar{x}_{h'}) = \sum_{j=1}^p (\bar{x}_{hj} - \bar{x}_{h'j})^2 . \quad (13)$$

The recalculation of the distance matrix is done as follows

$$D_{gg'} = \frac{1}{n_h + n_{h'}} \left(n_h \cdot D_{g'h} + n_{h'} \cdot D_{g'h'} - \frac{n_h \cdot n_{h'}}{n_h + n_{h'}} \cdot D_{hh'} \right) . \quad (14)$$

ard Method. The method uses a functional of the decomposition quality C_1 in formula (4). The criterion for the cluster connection is an increment to the total intra-group sum of the squares of observation deviations from the cluster average, thus

$$\Delta C_1 = \sum_{i=1}^g \sum_{j=1}^p (x_{gij} - \bar{x}_{gj})^2 - \sum_{i=1}^h \sum_{j=1}^p (x_{hij} - \bar{x}_{hj})^2 - \sum_{i=1}^{h'} \sum_{j=1}^p (x_{h'ij} - \bar{x}_{h'j})^2. \quad (15)$$

The increment is expressed as a sum of squares in an emerging cluster which is reduced by the sums of squares in both vanishing clusters. Using arithmetic modifications, the expression can be simplified into the form

$$\Delta C_1 = \frac{n_h \cdot n_{h'}}{n_h + n_{h'}} \cdot \sum_{j=1}^p (\bar{x}_{hj} - \bar{x}_{h'j})^2. \quad (16)$$

This equation is a product of the Euclidean distance between the centroids of clusters considered for the connection and a coefficient depending on the cluster size. The value of this coefficient grows with an increasing size of clusters, and for fixed $n_h + n_{h'}$ it represents the maximum in the case of same-size ($n_h = n_{h'}$) clusters. Since we create the connections to ensure the minimization of the criterion ΔC_1 , the Ward method tends to eliminate small clusters, i.e. to form those of roughly the same size, which is often a desirable property. Starting from the matrix of Euclidean distances between objects in the process of its modification, we can use the formula

$$D_{gg'} = \frac{1}{n_h + n_{h'} + n_{g'}} \cdot [(n_h + n_{g'}) \cdot D_{hg'} + (n_{h'} + n_{g'}) \cdot D_{h'g'} - n_{g'} \cdot D_{hh'}]. \quad (17)$$

The essence of this multidimensional statistical method is explained in detail in [7], [5] or [8]. Ward's method and Euclidean distance metric are the most used and have been also used in this analysis. Cluster analysis was based on data for 2015. There are various methods for determining the optimal number of clusters in cluster analysis, see for example [4]. However, there is no definitive answer to the question of determining the optimal number of clusters. The problem lies in that cluster analysis is basically an exploratory approach. Interpretation of the resulting hierarchical structure depends on the context, and there are often several solutions from the theoretical point of view.

3.2 Regression Analysis

The essence of regression and correlation analysis is explained in detail for example in [2] and this analysis was made on data for 2015, too. Average annual wage (only "average wage" in further text) was considered as dependent variable, real minimum wage (only "minimum wage" in further text), structural unemployment rate (only "unemployment rate" in further text) and GDP per head of person (only "GDP" in further text) were considered as independent variables.

The normality of all variables was verified in both ways, visually and using the Kolmogorov-Smirnov goodness of fit test. All variables were verified, Figure 1 and Table 2 present the results for variable „average wage”.

Although the variable "wage" is mostly lognormally distributed (i.e. with positive skewness), the variable "average wage" has a symmetrical distribution, which is in favour of the normal distribution, see Figure 1. P-value = 0.693242 shows that the hypothesis assumed normal distribution of average wage was not rebut at 5% significance level. Similar results were obtained from other variables.

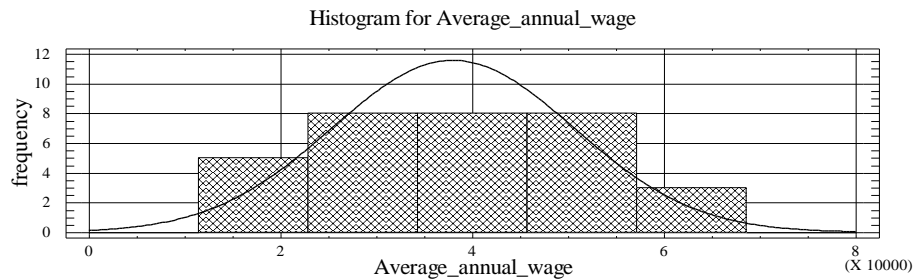


Fig. 1. Frequency histogram used for optical assessment of the normality distribution of average annual gross wage.

Table 2. Results for Kolmogorov-Smirnov goodness of fit test of normality for average annual gross wage.

<p>Estimated Kolmogorov statistic DPLUS = 0.125647 Estimated Kolmogorov statistic DMINUS = 0.115313 Estimated overall statistic DN = 0.125647 Approximate P-Value = 0.693242</p>

Table 3. Results for multiple linear regression analysis using the method of stepwise regression, forward selection (backward selection provides the same).

Multiple regression analysis					
Dependent variable: Average annual wage					
Parameter	Estimate	Standard error	T statistic	P-value	
CONSTANT	8609.67	3356.31	2.56522	0.0156	
GDP per head	0.76987	0.0826229	9.31788	0.0000	
Analysis of Variance					
Source	Sum of Squares	DF	Mean Square	F-Ratio	P-Value
Model	3.67012E9	1	3.67012E9	86.82	0.0000
Residual	1.26814E9	30	4.22713E7		
Total	4.93826E9	31			
R-squared = 74.3201 percent					
R-squared (adjusted for d. f.) = 73.4641 percent					
Standard Error of Est. = 6501.64					
Mean absolute error = 5323.46					
Durbin-Watson statistic = 1.62887					
Stepwise regression					
Method: forward selection					
F-to-enter: 4.0					
F-to-remove: 4.0					

Step 0:
0 variables in the model. 31 d. f. for error. R-squared = 0,00 % Adjusted R-squared = 0,00% MSE = 1.59299E8
Step 1:
Adding variable GDP per head of population with F-to-enter = 86.8228 1 variables in the model. 30 d. f. for error. R-squared = 74.32 % Adjusted R-squared = 73,46% MSE = 4.22713E7
Final model selected.

At the beginning, a linear regression hyperplane was considered. Stepwise regression with forward selection was used for determining a set of independent variables that have a significant effect on the dependent variable, see Table 3. It is clear from this table that only independent variable “GDP” was inserted into model, which narrows the model into straight line. Both individual t-tests and total F-test are significant. Durbin-Watson statistic is 1.62887, so in the interval (1,4; 2,6), which indicates that there is no problem with autocorrelation, and we can treat the residuals as independent. Determination index shows that 74.3201 percent of variability of average wage values is explain using linear regression model. Table 4 represents the results of quadratic regression function. Adjusted determination index of quadratic function is 87.2981 percent and that one of linear function is only 73.4641 percent. All individual t-tests and total F-test are significant at 5% significance level and the value of Durbin-Watson statistic 2.07188 shows that we have not a problem with autocorrelation. Thus, the polynomial regression function of the second degree better captures the dependence of “average wage” on “GDP”.

Table 4. Results for polynomial regression analysis.

Polynomial regression analysis					
Dependent variable: Average annual wage					
Parameter	Estimate	Standard error	T statistic	P-value	
CONSTANT	- 18289.0	5184.46	-3.52765	0.0014	
GDP per head	2.04375	0.226843	9.0095	0.0000	
GDP per head^2	- 0.000013	0.000002268	-5.80293	0.0000	
Analysis of Variance					
Source	Sum of Squares	D F	Mean Square	F-Ratio	P-Value
Model	4.35147E9	2	2.17574E9	107.53	0.0000
Residual	5.86783E8	29	2.02339E7		

Total (Corr.)	4.93826E9	31
R-squared = 88.1176 percent		
R-squared (adjusted for d.f.) = 87.2981 percent		
Standard Error of Est. = 4498.21		
Mean absolute error = 3612.72		
Durbin-Watson statistic = 2.07188		

Figures 2 and 3 represent the course of both types of dependencies considered and Figures 4 and 5 show the corresponding residual charts. It is clear from Figure 4 that residuals have not random character in the case of linear regression. In the case of polynomial regression, we can consider the course of residuals as satisfactory. In Addition to the visual approach, Glejser's test was used for heteroscedasticity testing in the case of polynomial regression. It was found on this basis that we have no problems with heteroscedasticity. For this reason, the polynomial regression function is more suitable model of dependence of "average wage" on "GDP". The sample regression parabola has the form

$$\text{Average wage} = -18,289.0 + 2.04375 * \text{GDP} - 0.000013 * \text{GDP}^2. \quad (18)$$

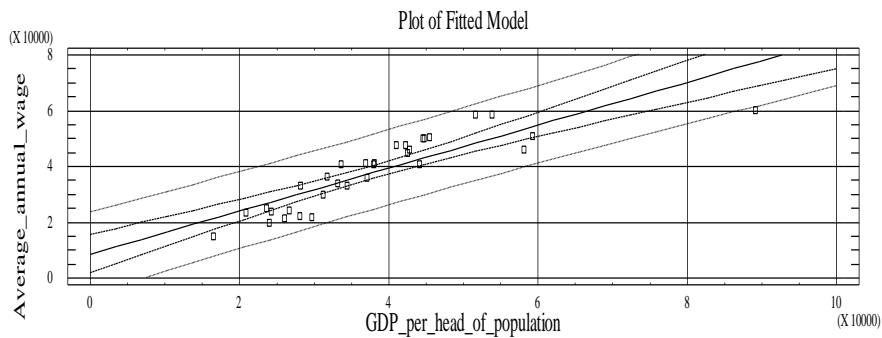


Fig. 2. Plot of fitted model – linear regression.

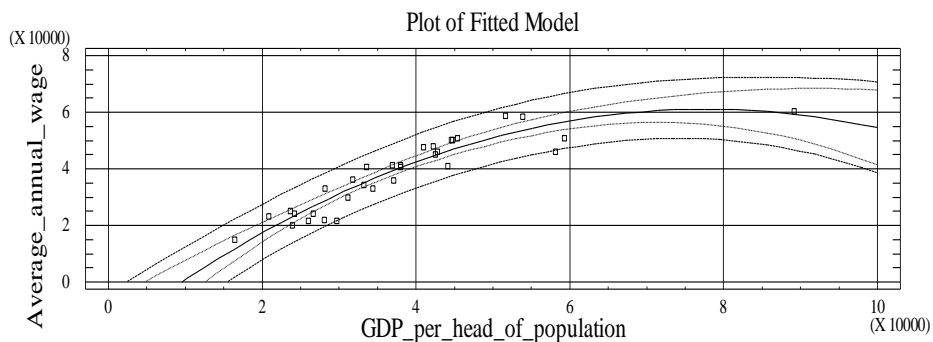


Fig. 3. Plot of fitted model – polynomial regression.

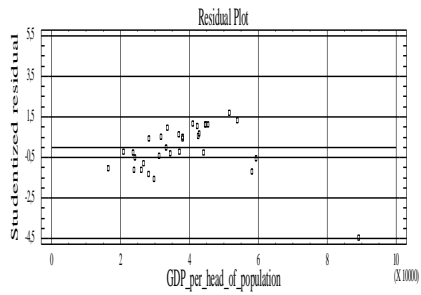


Fig. 4. Plot of residuals – linear regression.

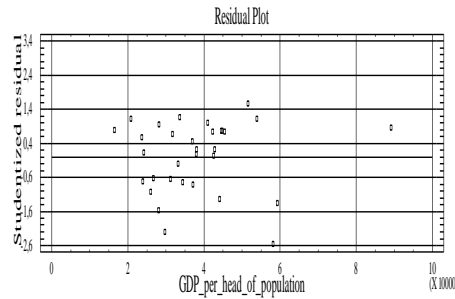


Fig. 5. Plot of residuals – polynomial regression.

4 Results and Discussion

Total 32 member states of OECD were chosen. There are covered three groups of states in total: non-European OECD member countries, developed Western European Countries and the former socialist bloc countries, see Table 1. Figures 6–9 presents the results of cluster analysis. Individual countries have been aggregated into five, seven, nine or eleven clusters to create clusters of the most similar countries in terms of average wage, minimum wage, unemployment rate and GDP in 2015.

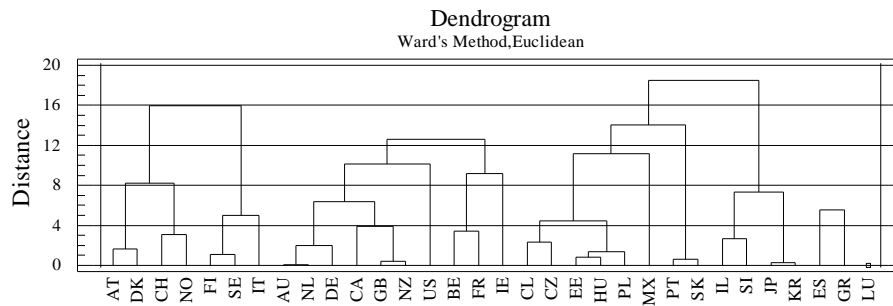


Fig. 6. Results of cluster analysis using Ward's method, Euclidean distance metric and five clusters.

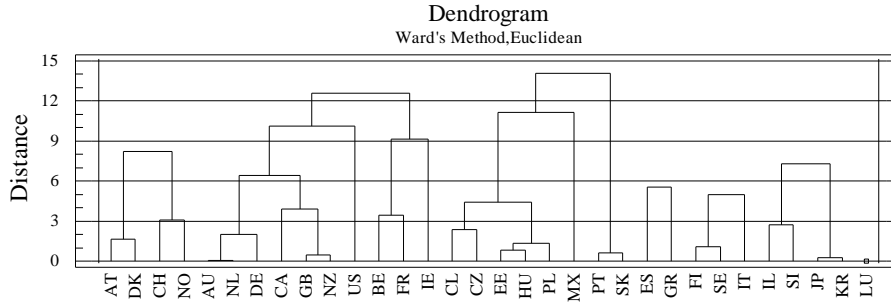


Fig. 7. Results of cluster analysis using Ward's method, Euclidean distance metric and seven clusters.

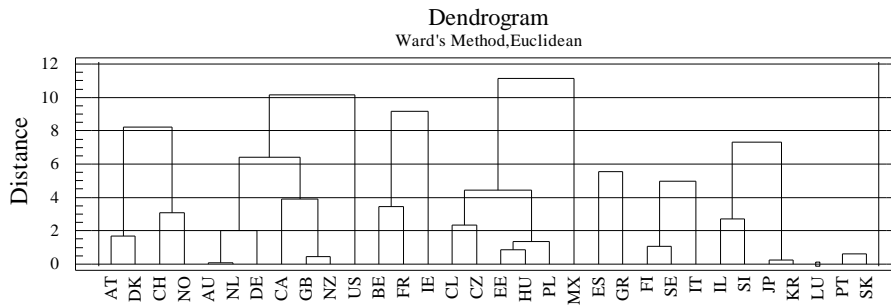


Fig. 8. Results of cluster analysis using Ward's method, Euclidean distance metric and nine clusters.

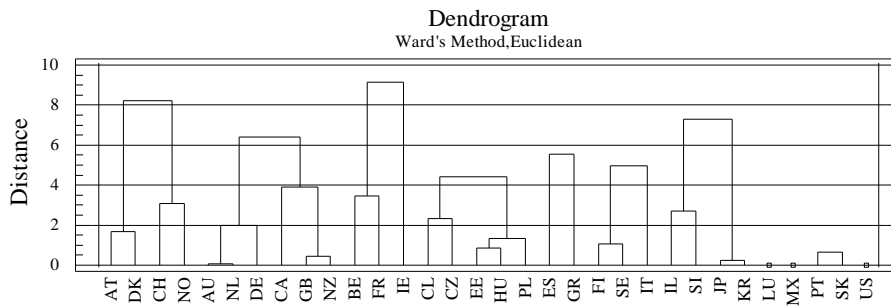


Fig. 9. Results of cluster analysis using Ward's method, Euclidean distance metric and eleven clusters.

When 32 selected countries are divided into five clusters in Figure 6, the first cluster is made up of seven countries, such as Austria, Switzerland, Denmark, Finland, Italy, Norway and Sweden. There are the Northern European countries and the most developed European countries. Another ten countries represent the second cluster: Australia, Belgium, Canada, Germany, France, Great Britain, Ireland, Netherlands,

New Zealand, and United States. There are high-developed European and non-European countries. The third cluster is made up of twelve countries. There are except for minor exceptions less developed non-European countries, states of the former socialist bloc and countries that solved substantial debt problems. There are Chile, Czech Republic, Estonia, Hungary, Israel, Japan, South Korea, Mexico, Poland, Portugal, Slovenia and Slovak Republic.

Table 5. Real and theoretical value of “average wage” in 2015 after conversion into purchasing power parity in USA calculated using regression function selected and variable “GDP”.

Country	Reality	Model	Country	Reality	Model
AU	50,167	47,157	IE	46,074	56,579
AT	46,084	45,368	IL	29,794	32,796
BE	47,702	43,629	IT	34,140	35,211
CA	47,843	44,872	JP	35,780	39,606
CH	58,389	54,075	KR	33,110	36,649
CL	23,247	18,619	LU	60,369	60,592
CZ	21,689	30,959	MX	14,867	11,809
DE	44,925	45,109	NL	50,670	47,718
DK	50,024	46,958	NO	50,908	57,178
EE	21,564	26,098	NZ	39,888	35,708
ES	36,325	33,466	PL	23,998	23,568
FI	40,731	40,573	PT	24,105	26,968
FR	41,252	39,455	SE	40,909	46,592
GB	41,384	40,640	SI	33,085	29,009
GR	25,211	22,783	SK	22,031	28,869
HU	19,999	23,183	US	58,714	52,550

The fourth cluster has only two members Spain and Greece, which solved extensive debt problems recently. Only one state Luxembourg with absolutely the highest average wage represents the fifth cluster. Figures 7–9 show division of 32 selected countries into seven, nine or eleven clusters.

From the results of regression and correlation analysis made for 2015 is clear that only variable “GDP” from considered three independent variables influences statistically significant dependent variable “average wage”. Determination index acquires the value 88,12 %. It means that 88.12 % of variability of observed values of “average wage” contrived to explain using selected quadratic regression function and “GDP” variable. The concave parabola with the maximum for 157,212 USD PPP of “GDP” represents a regression function describing the dependency of “average wage” on “GDP”. It means that “average wage” increases with increasing “GDP” as far as 157,212 USD PPP. As soon as this point is reached, the “average wage” would start to decline with “GDP” growth. On the other hand, none of the countries is far below such a high “GDP”. Table 5 represents the real and theoretical values of “average wage” calculated using “GDP” in 2015.

5 Conclusion

The highest average wages are in the most economically advanced countries in the world. The average wage represents a criterion of financial prosperity of the country. Average wage after its conversion into purchasing power parity reflects different living costs in individual countries. Wage distribution is positively skewed and so, the most of people do not reach the average wage.

Income differences between individual OECD member countries are lower taking into account the prices of goods and services than in the case of the comparison of nominal average wages. Even in purchasing power parity, the highest average wage is in Luxembourg, United States and Switzerland. On the contrary, we can see the lowest average wage in Mexico, and Hungary. The highest costs of living are also in the OECD member countries with the highest average wage. Especially, expenditures on housing and services are considerably higher in Luxembourg, United States, Switzerland Norway or Germany than for example Mexico, Chile or Hungary. For example, while the average wage in Switzerland is higher than in Mexico 15.4 times when comparing the absolute values, the average wage in purchasing power parity is only 5.4 times. However, the gap has been steadily risen in recent years. When comparing gross wages in absolute values, Luxembourg's, United States's, Switzerland's or Norway's financial advantage is very high, in purchasing power parity this is not true so much, because these countries have also the highest prices for goods and services in the OECD member countries.

The lowest wage differences are in the Czech and Slovak Republics and in the Scandinavian countries. This means, among other things, more employees reach average and higher wage than in other countries. The highest wage differences are in the Mexico. Lower wage differences and a functional social network are the reasons, why the least poor OECD citizens are in the Czech Republic. The poor is considered a citizen with an income of less than 60 % of the median wage. For households, the income is calculated for determining the poverty line.

In the European OECD member countries, employees pay more for income tax and compulsory insurance than in the non-European countries. From the financial point of view, it would be the best to get an average wage for example in Switzerland, to pay a tax in Chile and to spend a net wage on the purchase of goods and services in Mexico.

The main research hypothesis (division into Western European countries on the one hand and Eastern European countries on the other hand is still holding) can be considered proved.

Acknowledgements. This paper was subsidized by the funds of institutional support of a long-term conceptual advancement of science and research number IP400040 at the Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic.

References

1. Bartošová, J., Želinský, T.: Extent of Poverty in the Czech and Slovak Republics 15 Years after the Split. *Post-Communist Economics* 25(1), 119–131 (2013), DOI: 10.1080/14631377.2013.756704.
2. Darlington, R. B., Hayes, A. F.: *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*. The Guilford Press, New York (2017).
3. Doucek, P., Nedomová, L.: How the Grade of Education Influences Wages in Different Phases of Economic Cycle: Case the Czech Economy. In: *International Conference on Education, E-learning and Management Technology (EEMT 2016)* [online]. Xi-an, 27.08.2016–28.08.2016. Čína: Atlantis–Press, pp. 49–55 (2016).
4. Löster, T.: Evaluation of Coefficients for Determining the Optimal Number of Clusters in Cluster Analysis on Real Data Sets. In: *The 9th International Days of Statistics and Economics (MSED 2015)* [online]. Praha, 10.09.2015–12.09.2015. Slaný: Melandrium, pp. 1014–1023 (2015).
5. Makhlova, E., Pecáková, I.: The Fuzzy Clustering Problems and Possible Solutions. In: *The 9th International Days of Statistics and Economics (MSED 2015)* [online]. Praha, 10.09.2015–12.09.2015. Slaný: Melandrium, pp. 1052–1061 (2015).
6. Malá, I.: Multivariate Probability Model for Incomes of the Czech Households. *Politická ekonomie* 63(7), 895–908 (2015), DOI: <https://doi.org/10.18267/j.polek.1040>.
7. Rencher, A. C., Christensen, W. F.: *Methods of Multivariate Analysis*. John Wiley & Sons, New Jersey (2012).
8. Šimpach, O.: Application of Cluster Analysis on the Demographic Development of Municipalities in the Districts of Liberecký Region. In: *The 7th International Days of Statistics and Economics* [online]. Prague, 19.09.2013–21.09.2013. Slaný: Melandrium, pp. 1390–1399 (2013).
9. OECD Statistics, <https://stats.oecd.org/>.