

Models and Forecasts of Wage Distributions in Czech Education and Health Care Sectors

Diana BÍLKOVÁ

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

Abstract. This paper deals with the development of wages in Czech education and health care sectors since the beginning of the global economic crisis to the present. Data for this research come from the official website of the Czech Statistical Office and they include employees of the Czech Republic. The variable surveyed is the gross nominal monthly wage of the employee. The construction of wage models in Czech education and health care sectors that allow some comparison between education and health care sectors, is an important aim of this research. Three-parameter lognormal curves are the basis of these models. The construction of predictions not only the wage levels in the Czech education and health care sectors until 2020, but also of the predictions of the whole wage distributions in these two sectors until 2020 is a part of this study. Important aim is the construction of clusters of sectors that are the most similar according to average wage and selected variables.

Keywords: Wage Models, Lognormal Curves, Maximum Likelihood Method, Wage Predictions, Cluster Analysis.

1 Introduction

In general, employees performing psychologically demanding work achieve the highest wage levels, while employees performing physically demanding work achieve lower wage levels, and the highest monthly wage fluctuations are monitored for managers, whereas the smallest wage differences in monthly wages are detected for office workers and teachers, which are not very different throughout the whole Czech Republic.

Through the whole society, the opinion eternally prevails that education of the youngest generations is among the priorities, and work in education therefore deserves a solid reward.

For many years, governments in the Czech Republic have relied on the fact that if someone graduates from a Medical Faculty or Medical School, they will devote their profession even under a relatively small reward. Especially, if poetic slogans about the noble profession and help to sick are added to it. After all, this is not even work, but this is a mission. If someone decides to join his (her) professional life with health care already in youth and successfully passes through the pitfalls of the appropriate

educational institutions, it is very likely that they will remain in health care in decent working conditions. However, non-health workers are necessary for the reliable functioning of hospitals, too. Drivers, cooks, caretakers, helpers, hospital attendants and many other professions. And the labour market hand is starting to threaten here.

This paper focuses on research of wage development in Czech education and health care sectors since the beginning of the global economic crisis to the present. The aim is to capture the predictions until 2020, not only predictions of wage level, but also predictions of the whole wage distributions, since it is necessary to know the full wage distribution of the population from the point of view of a correct evaluation of the wage aspect of the standard of living of the population and within correct decisions in terms of steps in this area. Important aim is the construction of clusters of sectors that are the most similar according to average wage and selected variables in 2017.

Although the beginning of the global economic crisis can be dated to the autumn of 2008, so the consequences of its accession were economically manifested especially in 2009. For this reason, the study covers the period 2009–2017 with predictions for the period 2018–2020.

2 Data Base

Data for this research come from the official website of the Czech Statistical Office (CSO). There are annual data in the form of an interval frequency distribution with extreme open intervals. The researched variable is the gross (nominal) monthly wage in CZK. Employees of the Czech Republic represent the statistical units surveyed.

Table 1. Exact and abbreviated (modified) names of the sectors analysed [1].

Exact name	Modified name
Education	Education system
Human health and social work activities	Health service
Financial and insurance activities	Finance and insurance
Information and communication	Informatics and communications
Accommodation and food service activities	Accommodation activities
Administrative and support service activities	Administrative activities

1

Data from the CSO include employees in the business and non-business sphere in the Czech Republic. The wage is paid to the employee for the work done in the private (business) sphere, the salary in the budget (state, public, non-business) sector. In general, salaries in the non-business sphere are much more nivelised than wages in the business sphere, where very high wage variability exists. In terms of data presented on the CSO website, wages in the business sector and salaries in the non-business sector are below the wage term.

Abbreviated (modified) names of the analysed sectors are used in the following text (figures and tables). The exact names according to the official CSO website and abbreviated (modified) names of the analysed sectors are presented in Table 1.

3 Theory and Methods

Wage distribution models are based on the construction of three-parameter lognormal curves, the parameters of which are estimated using maximum likelihood method. Three-parameter lognormal distribution is defined with probability density function

$$f(x; \mu, \sigma^2, \theta) = \frac{1}{\sigma \cdot (x - \theta) \cdot \sqrt{2\pi}} \cdot \exp\left[-\frac{[\ln(x - \theta) - \mu]^2}{2\sigma^2}\right], \quad x > \theta, \quad (1)$$

$$= 0, \quad \text{otherwise.}$$

The essence of three-parameter lognormal curves used in modelling wage distributions is explained in [3] or [5] and the essence of the maximum likelihood method used for point parameter estimation of these lognormal curves is explained in [4]. For a given parameter θ (minimum wage), the estimates of the remaining two parameters of three-parameter lognormal curves are taken using

$$\hat{\mu}(\theta) = \frac{\sum_{i=1}^n \ln(x_i - \theta)}{n}, \quad (2)$$

$$\hat{\sigma}^2(\theta) = \frac{\sum_{i=1}^n [\ln(x_i - \theta) - \hat{\mu}(\theta)]^2}{n}. \quad (3)$$

The essence of time series analysis and their simple characteristics are explained in [6]. The predictions of wage level (average and median-middle gross monthly wage) by sectors until 2020 were created on basis of the respective time series from the period 2009–2017. In the context of trend development, exponential smoothing was applied in time series analysis. Exponential smoothing is one of the adaptive approaches to modelling time series and it uses the weighted least square method, where scales exponentially decreasing towards the past. The advantage of exponential smoothing lies in the fact that the most recent observations have the highest weight. Appropriate exponential smoothing was chosen using interpolation criteria. The statistical software automatically evaluates the most advantageous combinations of equalizing constants α and β . Sample residual autocorrelation functions and sample residual partial autocorrelation functions show that the non-systematic component does not show autocorrelation in all cases, and consequently the relevant exponential smoothing is satisfactory. The Durbin-Watson statistics are close to two in all cases,

i.e. always in the interval (1.6, 2.4). Random faults can be therefore considered as independent. It can be approached to using Theil coefficient of mismatch to evaluate the model's quality. The annual time series are abbreviated of m observations (in this case, $m = 3$ observations), with forecasts for these $m = 3$ years being made using the corresponding exponential smoothing. Theil coefficient of mismatch gets the low zero boundary only in case of flawless forecasts. The more the Theil coefficient of mismatch deviates from zero, the more the prediction differs from ideal flawless prognoses. The square root of the Theil mismatch coefficient can be interpreted as a relative prediction error. All calculated values of Theil mismatch coefficient and relative prediction error indicate the high quality of the selected exponential smoothing models. The essence of cluster analysis is explained in detail for example in [2], and Ward's method and Euclidean distance were used, since there are the most widely used techniques within cluster analysis, and number of clusters was determined five using Dunn index.

4 Results and Discussion

Figures 1–4 offer an idea of the development of the entire wage distribution of the health care sector and the education sector over time, and they allow a comparison of the wage distribution shape for these two sectors over time. There are model wage distributions constructed using three-parameter lognormal curves, since empirical wage distributions could not be used for the reason of unequally wide intervals of empirical distributions with extreme open intervals. The values in Table 2 were calculated from these model distributions.

From Table 2, the shift in wage distributions is evident in both sectors between the period of the global economic crisis and the present. For example, while in 2009, 16.8 % of employees in the education sector reached a maximum wage of 15,000 CZK, in 2014, after the crisis, this share fell to 13.7 % and in 2017 even to 4.4 %. In health care sector, in 2009, 21.2 % of employees reached a maximum wage 15,000 CZK, in 2014, after the crisis, this share fell to 17.1 % and in 2017 even to 4.9 %.

It is evident from Table 2, that approximately 64.8 % of employees in education sector and 63.7 % of employees in health care sector had their gross monthly wage lower or at most equal to the countrywide average gross monthly wage in 2017. This means that approximately two-thirds of employees in both analysed sectors did not achieve average wage in 2017. On the contrary, the share of employees whose gross monthly wage exceeded CZK 100,000 does not exceed 1 % over the whole analysed period for both sectors. Therefore, we do not see any significant changes at higher wage limits throughout the whole period.

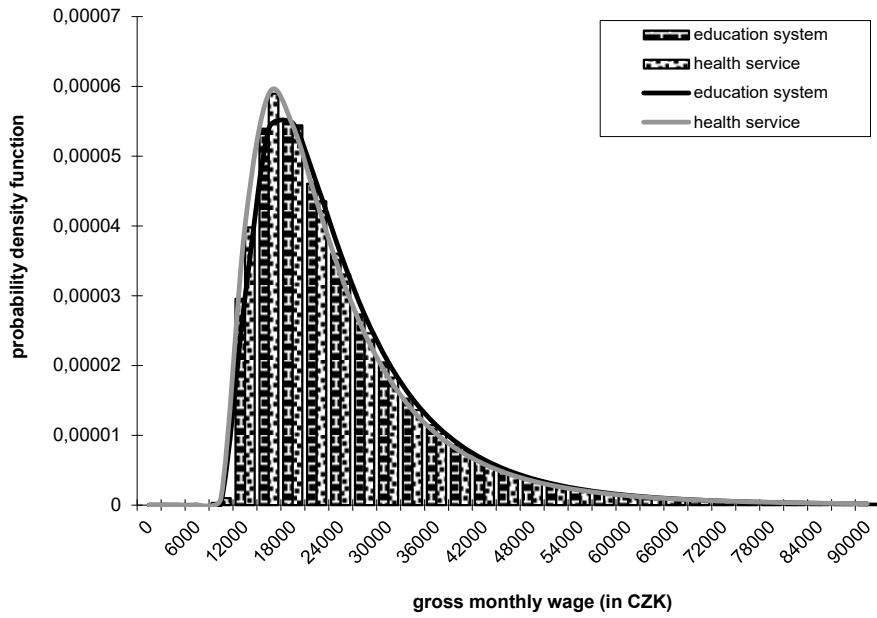


Fig. 1. Model distribution of gross monthly wages in the education and health sectors in 2009.

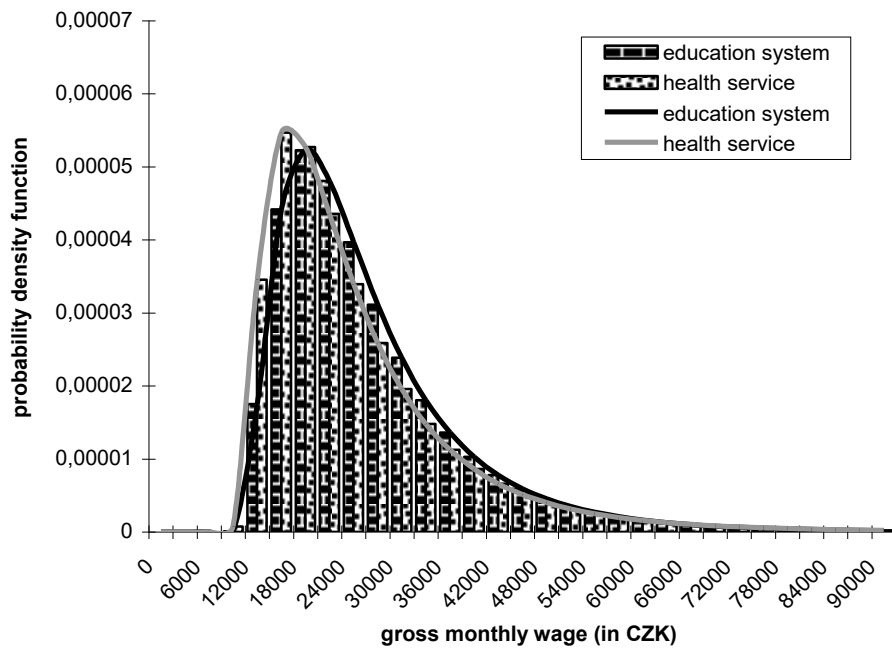


Fig. 2. Model distribution of gross monthly wages in the education and health sectors in 2011.

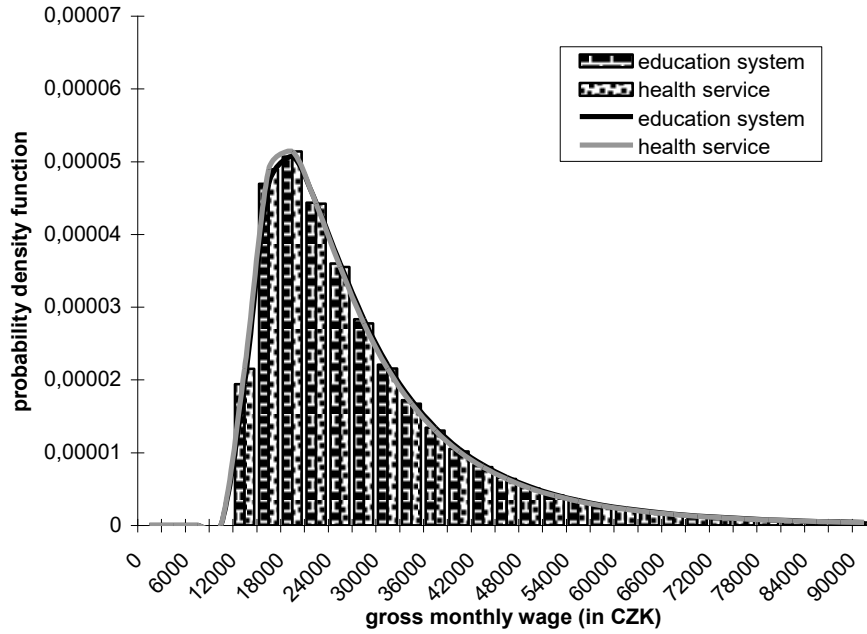


Fig. 3. Model distribution of gross monthly wages in the education and health sectors in 2015.

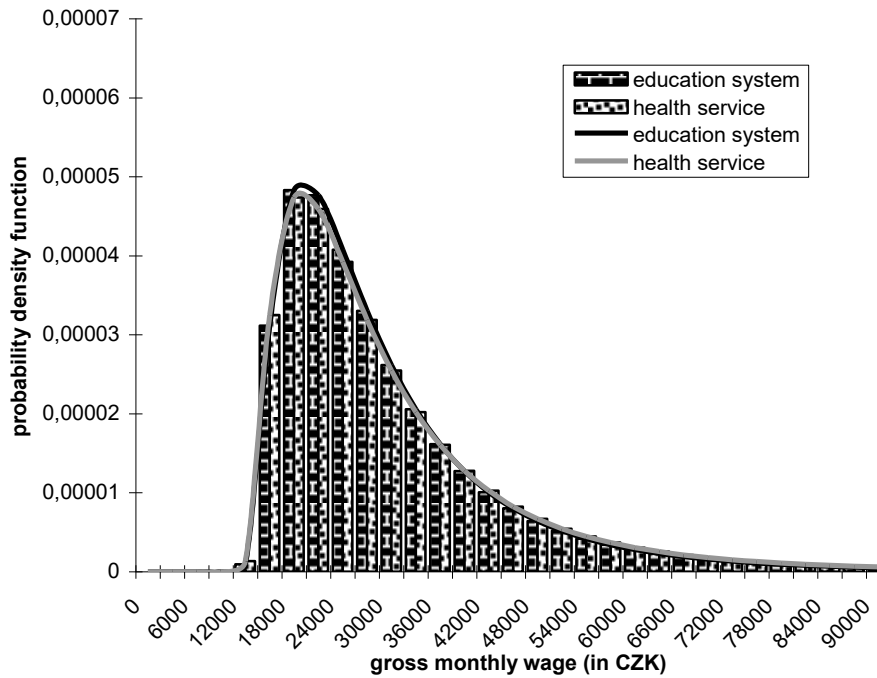


Fig. 4. Model distribution of gross monthly wages in the education and health sectors in 2017.

Table 2. Estimated proportions (in percentages) of employees whose gross monthly wage reaches at most upper wage limit in 2009, 2014 and 2017.

Upper limit	Education system			Health service		
	2009	2014	2017	2009	2014	2017
15,000	16.8	13.7	4.4	21.2	17.1	4.9
20,000	44.0	39.0	26.7	48.7	44.0	27.0
30,000	77.6	73.1	64.8	79.7	76.5	63.7
40,000	90.7	87.7	82.8	91.4	89.6	81.5
50,000	95.8	94.0	91.0	96.0	95.0	89.9
60,000	98.0	96.9	95.0	98.0	97.4	94.2
70,000	99.0	98.3	97.1	98.9	98.6	96.5
80,000	99.4	99.0	98.2	99.4	99.2	97.8
90,000	99.7	99.4	98.9	99.7	99.5	98.6
100,000	99.8	99.6	99.3	99.8	99.7	99.0

Figures 5 a 6 demonstrate illustrations of exponential smoothing of time series of average wage for education and health care sectors, where Brown's quadratic exponential smoothing was chosen as the most appropriate in time series of education sector and Holt's linear exponential smoothing was chosen as the most appropriate in time series of health care sector. This was done analogically in terms of exponential smoothing for median wage, wage standard deviation (necessary for the construction of the prediction of the whole wage distributions until 2020) in both sectors and national average wage (necessary for construction of predictions of national average wage until 2020). The value of 32,456 was used for the estimation of minimum wage for 2020.

Based on this trend analysis of the corresponding time series, the predictions of average and median gross monthly wages for both analysed sectors were constructed, see Table 3. We can see from this table that we can expect gradual growth of wage levels in both sectors until 2020, where growth of wage level is expected to be faster in health care sector (on average more than 3 % per annum) than in education sector (on average about 2 % per annum).

The minimum wage in 2018 was 12,200 CZK. The government's planned valorisation of the minimum wage for 2019 is CZK 13,200. The government of the Czech Republic expects that the minimum wage could be regularly valorised from 2020 onwards. This would be based on the average wage and set as its 0.44 times. For this reason, the assumed minimum wage for 2020 was estimated to be 0.44 times the projected national average wage (Table 4) in this year

$$32,456 * 0.44 = 14,280.64 = 14,300 \text{ CZK.} \quad (4)$$

Table 3. Prediction of wage level development (in CZK) and wage variability development (standard deviation in CZK, coefficient of variation in percentages) to 2020 for education and health care sectors.

Year	Education sector				Health care sector			
	Average	Median	Std. d.	Var. c.	Average	Median	Std. d.	Var. c.
2018	29,778	28,306	14,881	49.97	32,816	28,031	20,278	61.79
2019	30,771	29,217	15,056	48.93	33,926	28,965	20,981	61.84
2020	31,862	30,216	15,230	47.80	35,036	29,898	21,685	61.89

Table 4. Prediction of national average wage development (in CZK) to 2020.

Year		
2018	2019	2020
30,110	31,283	32,456

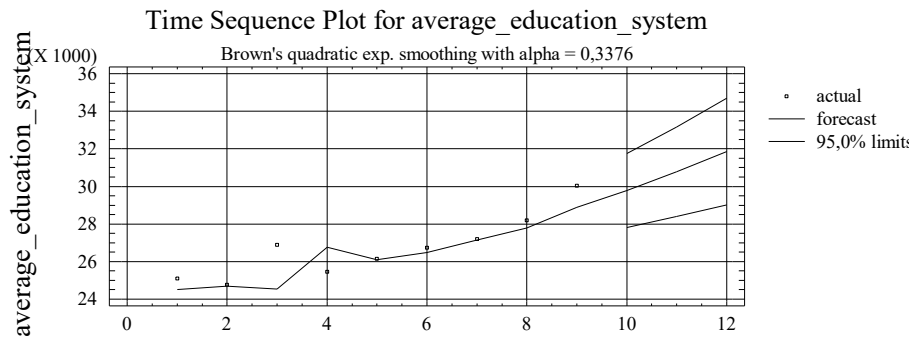


Fig. 5. Brown's quadratic exponential smoothing – average of education sector.

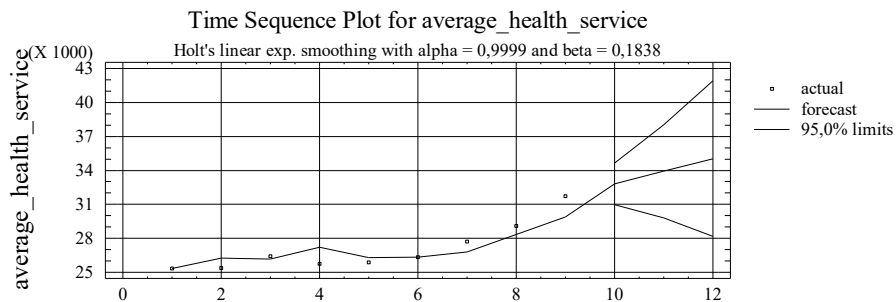


Fig. 6. Holt's linear exponential smoothing – average of health care sector.

Figures 7–9 represent predictions of the full wage distributions for the education and health care sectors until 2020. Middles of wage intervals with wide of 3,000 CZK are on the vertical axis of Fig. 7–9.

doi: 10.36689/uhk/hed/2019-01-004

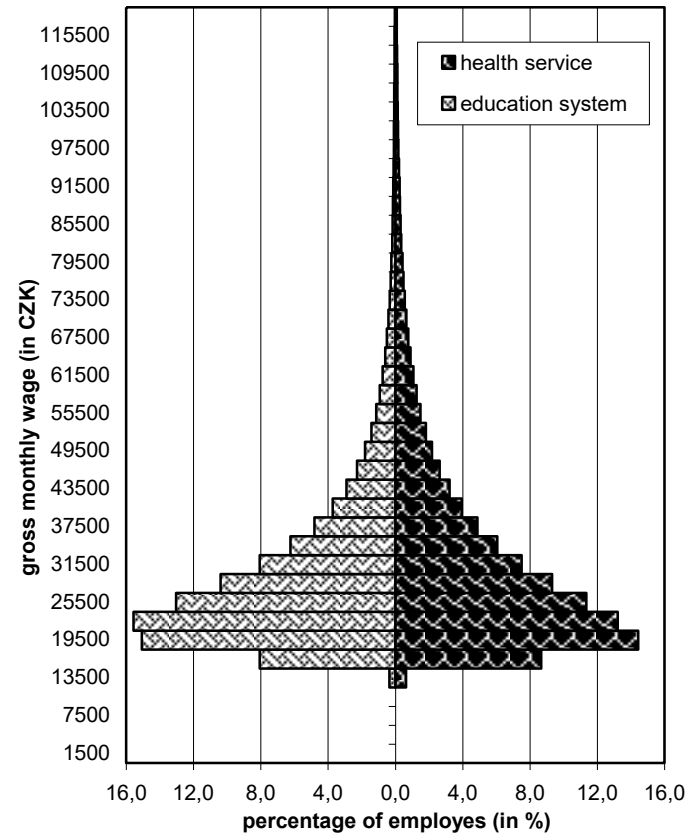
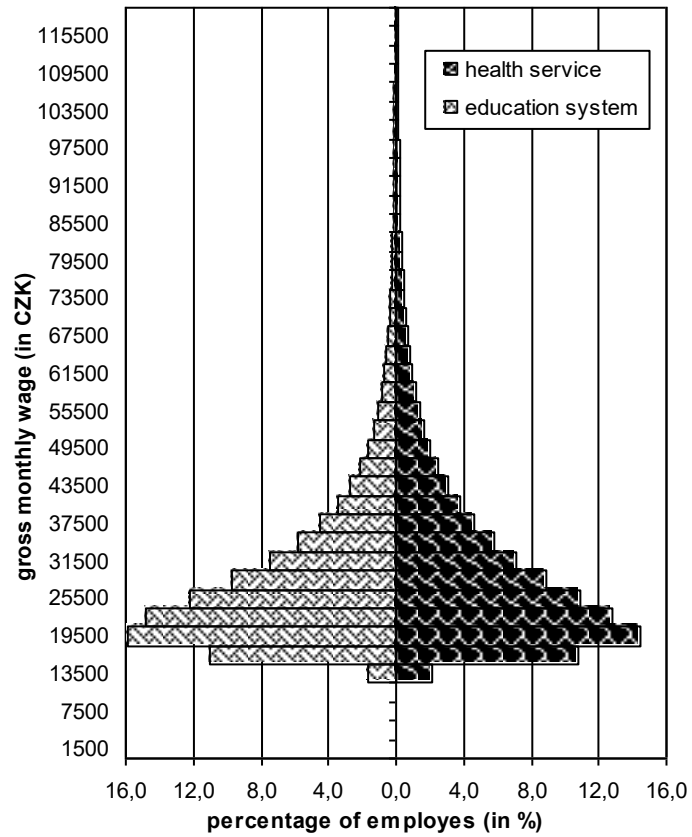


Fig. 7. Prediction of the distribution of the gross monthly wage (in percentages) for education and health care sectors for the year 2018.

Fig. 8. Prediction of the distribution of the gross monthly wage (in percentages) for education and health care sectors for the year 2019.

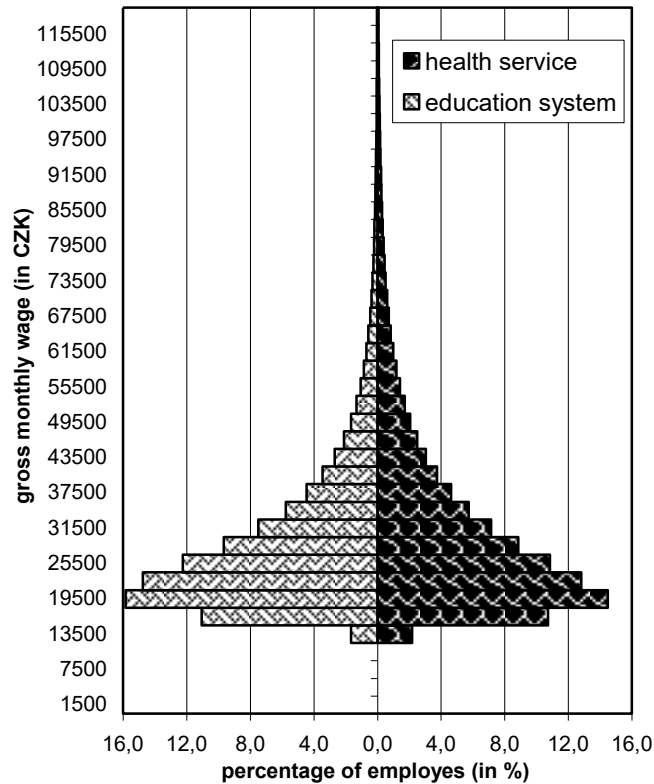


Fig. 9. Prediction of the distribution of the gross monthly wage (in percentages) for education and health care sectors for the year 2020.

We can see from these figures a higher percentage representation of lower wage intervals for the education sector compared to the health care sector and a higher percentage representation of higher wage intervals for health care sector compared to the education sector. This situation is probably related to the expected higher wage level in the health care sector than in the education sector over the next three years. The zero percentage representation at the lowest wage intervals is related to the existence of a minimum wage institute and its gradual increase.

Figure 10 shows the results of cluster analysis applied to all sectors of the Czech economy based on cross-sectional data. We can see from this figure that the first cluster has four members: 1. Agriculture, forestry and fishing; 2. Construction; 3. Real estate activities; 4. Arts, entertainment and recreation. The second cluster is five-membered: 1. Industry; 2. Wholesale and retail trade; repair of motor vehicles and motorcycles; 3. Transportation and storage; 4. Education; 5. Human health and social work activities. This means that the two sectors analysed in this study are in the same cluster. The third cluster contains only two the worst-paid sectors: 1. Accommodation

and food service activities; 2. Administrative and support service activities. On the contrary, the fourth cluster contains two the best-paid sectors: 1. Information and communication; 2. Financial and insurance activities. The fifth cluster has the two remaining clusters: 1. Professional, scientific and technical activities; 2. Public administration and defence; compulsory social security.

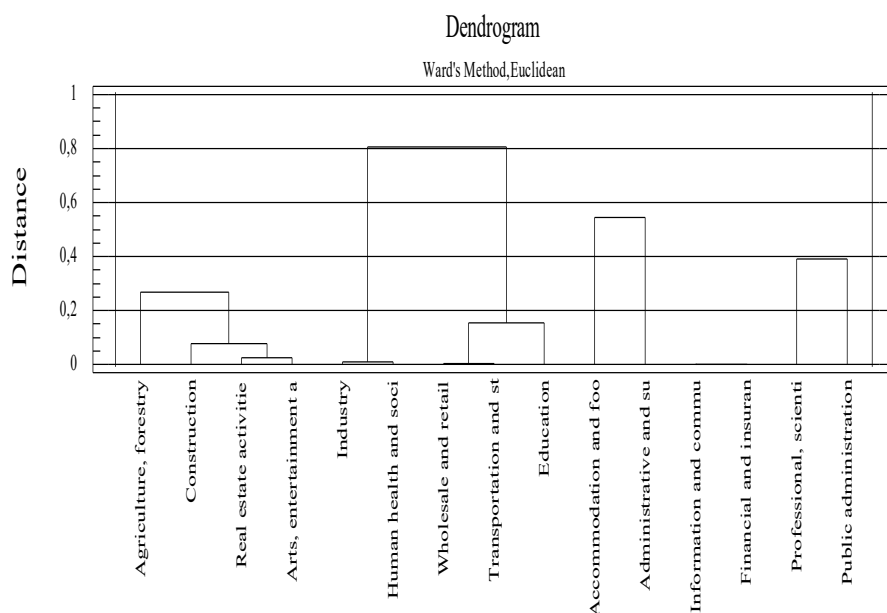


Fig. 10. Dendrogram of the results of cluster analysis: Clusters of all sectors of the Czech economy that are the most similar according average wage and related indicators in 2017.

5 Conclusion

The big problem in these sectors is probably the above-average high tax on labour, which drains from the net income of the people, and at the same time it demotivates employers from adding to employees. There is no way in increase of the minimum wage. If we want in order that people will take more money, we need to increase the productivity. At first people must get rich, otherwise the country is not rich. The recipe may consist in a reduction in load levies. At present, a very good situation is for people who want to change their job title. Wages in the Czech Republic are considerably underestimated over other European Union countries. Trade unions play an important role here as they help to exert pressure to increase wages.

Wages in the health care sector are slightly more differentiated, while wage diversification is roughly the same in both sectors, and for the future, greater wage diversification is expected for the health care sector than for the education sector. At present, about two-thirds of employees in the education and health care sectors have

a lower their gross monthly wage than the national average gross monthly wage in the Czech Republic. We can expect a gradual increase in wage levels in the education and health sectors until 2020. The average gross monthly wage in these two sectors is approximately the same as in sectors of industry, wholesale and retail trade, repair of motor vehicles and motorcycles and transportation and storage.

Now it is necessary to increase nurses' tariffs. It is important that young doctors and qualified nurses do not leave abroad immediately after school. The biggest problem is the lack of nurses at present. Systemic arrangements, such as tariff settlements or surcharges for work shifts of inpatient care, have not been addressed in the last electoral period. Health care and education sectors would be given a priority in terms of wage increases. In the Czech Republic, this would also strengthen the role of primary care because hundreds medical practitioners lack. In addition, the average age of medical practitioners ranges from 58 to 59 years.

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. Český statistický úřad, <http://www.czso.cz>, last accessed 2018/11/21.
2. Everitt, B. S., Landau, S., Leese, M., Stahl, D.: Cluster Analysis. 5th edn. John Wiley & Sons Ltd., West Sussex (2011).
3. Johnson, N. L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, Vol. 1. 2nd edn. Wiley-Interscience, New York (1994).
4. Johnson, N. L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, Vol. 2. 2nd edn. Wiley-Interscience, New York (1995).
5. Kleiber, C., Kotz, S.: Statistical Size Distributions in Economics and Actuarial Sciences. 1st edn. Wiley-Interscience, New York (2003).
6. Shumway, R. H., Stoffer, D. S.: Time Series Analysis and Its Applications: With R Examples. 4th edn. Springer, New York (2017).