THE WORST PAID SECTORS OF THE CZECH ECONOMY

Diana Bílková

Abstract

This paper deals with the issue of the worst paid sectors of the Czech economy, which are Accommodation and Food Service Activities and the Administrative and Support Service Activities sectors. There is a development of model distribution of monthly earnings (wages and salaries together) of employees in individual years 2009–2020. Three-parameter lognormal curves were used as the basic theoretical probability distribution of earning models. The method of L-moments of point parameter estimation is used to estimate the parameters of lognormal curves, while the beginning of lognormal curves is estimated by a minimum wage. The data come from the official website of the Czech Statistical Office. Based on the models of earning distributions, the percentage shares of employees in individual earning intervals of the same length of five thousand crowns were estimated. The results show higher percentages of employees in lower earning intervals for Accommodation and Food Service Activities sector compared to Administrative and Support Service Activities sector. The percentages of employees in higher earning intervals are lower for Accommodation and Food Service Activities sector than for Administrative and Support Service Activities sector.

Key words: The worst paid sectors and earning distribution, three-parameter lognormal curve and L-moment method, goodness-of-fit test

JEL Code: E24, C55, C51

Introduction

The method of L-moments of point estimation of parameters (Hosking 1990; Hosking, 2006) has been used in the past, for example, in connection with the study of extreme precipitation (Kyselý & Picek, 2007) or monthly precipitation (Guttman, 1994). Studies (Pearson, 1991; Kumar & Chatterjee, 2005; Eslamian & Feizi, 2007; Noto & La Loggia, 2009; Saf, 2009; Alahmadi, 2017) examine the issue of L-moments in connection with flood frequency and rainfall frequency analyses.

This paper deals with the application of the L-moment method to economic data from the labour market. The L-moment method was used for point estimation of parameters of three-parameter lognormal curves representing models of distribution of monthly earnings (wages and salaries together) of employees in two sectors of the Czech economy with the lowest earnings in the period 2009‒2020. There are Accommodation and Food Service Activities and Administrative and Support Service Activities sectors.

The beginning of these curves was estimated using the minimum wage in the corresponding year. The development of the minimum wage amounts in the period under review is shown in Table 1. However, the question of the suitability of a given curve for the earning distribution model is not a completely common mathematical-statistical problem, in which we test the null hypothesis “H_0: Random sample comes from the assumed theoretical distribution” against the alternative hypothesis “H_1: non H_0”, since in the case of the distribution of earnings, typical samples are huge (in this case tens to hundreds of thousands of respondents), therefore the goodness-of-fit test always leads to the rejection of the null hypothesis about the assumed distribution. This is due not only to the fact that at such large sample sizes, the power of the test is so strong that the goodness-of-fit test reveals all the slightest deviations of the actual distribution of earnings and model, but also from the principle of the test. However, we are not practically interested in small deviations, so only the approximate conformity of the model with reality suffices. When evaluating the suitability of the model, it is necessary to proceed to a large extent subjectively and rely on logical analysis and experience. For this reason, goodness-of-fit tests are not performed in this study.

Tab. 1: Development of the minimum wage amounts in 2009‒2020[1]

Year	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
Minimum wage amount	8,000	8,000	8,000	8,000	8,208	8,500
Year	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Minimum wage amount	9,200	9,900	11,000	12,200	13,350	14,600

Source: www.mpsv.cz

The main aim of this research is to present the application of the L-moment method to economic data, specifically to data from the labour market. Another objective is to compare the development of the distribution of earnings of two sectors of the Czech economy with the lowest level of earnings.

The data come from the official website of the Czech Statistical Office. These were data in the form of interval frequency distribution with unequally wide earning intervals and

[1] In 2013, the minimum wage was CZK 8,000 in the period from the 1st January to the 31st July and CZK 8,500 in the period from the 1st August to the 31st December. The amount of the minimum wage in Table 1 represents the average minimum wage in 2013.

with extreme open intervals. The data were processed using a Microsoft Excel spreadsheet and the SPSS statistical programming environment.

1 L-moments

L-moments represent an alternative system describing the shape of the probability distribution. They are an analogy of conventional moments, but they can be estimated on the basis of a linear combination of order statistics, i. e. L-statistics.

Let X be a random variable having a continuous distribution with a distribution function $F(x)$ and a quantile function $x(F)$. Let $X_1, X_2, \ldots, X_n$ be a random sample of range n from this distribution. Then $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ are the order statistics of random sample of range n, which comes from the distribution of a random variable X.

1.1 L-moments of probability distribution

L-moment of the r-th order of the random variable X is defined

$$\lambda_r = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot E(X_{r-j:r}), \quad r = 1, 2, \ldots, \tag{1}$$

the expected value of the order statistic has the form

$$E(X_{r:n}) = \frac{n!}{(r-1)! \cdot (n-r)!} \cdot \int_0^1 x \cdot [F(x)]^{r-1} \cdot [1 - F(x)]^{n-r} \mathrm{d} F(x). \tag{2}$$

The first four L-moments of the probability distribution are now defined

$$\lambda_1 = E(X_{1:1}) = \int_0^1 x(F) \, \mathrm{d} F, \tag{3}$$

$$\lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2}) = \int_0^1 x(F) \cdot (2F - 1) \, \mathrm{d} F, \tag{4}$$

$$\lambda_3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3}) = \int_0^1 x(F) \cdot (6F^2 - 6F + 1) \, \mathrm{d} F, \tag{5}$$

$$\lambda_4 = \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \int_0^1 x(F) \cdot (20F^3 - 30F^2 + 12F - 1) \, \mathrm{d} F. \tag{6}$$

The probability distribution can be specified by its L-moments even if some of its conventional moments do not exist, but the opposite is not true. It is often appropriate to standardize higher L-moments λ_r, $r \geq 3$ to be independent of the unit of measure of random variable X. The ratio of L-moments of random variable X is defined

$$\tau_r = \frac{\lambda_r}{\lambda_2}, \quad r = 3, 4, \dots. \tag{7}$$

Using equations (3)–(5) and equation (7), we obtain equations for the case of a three-parameter lognormal distribution[4]

$$\lambda_1 = \theta + \exp\left(\mu + \frac{\sigma^2}{2}\right), \tag{8}$$

$$\lambda_2 = \exp\left(\mu + \frac{\sigma^2}{2}\right) \cdot \operatorname{erf}\left(\frac{\sigma}{2}\right), \tag{9}$$

$$\tau_3 = \frac{6}{\sqrt{\pi} \cdot \operatorname{erf}\left(\frac{\sigma}{2}\right)} \cdot \int_0^{\sigma/2} \operatorname{erf}\left(\frac{x}{\sqrt{3}}\right) \cdot \exp(-x^2)\, \mathrm{d}x. \tag{10}$$

1.2 Sample L-Moments

Let $x_1, x_2, \dots, x_n$ be a random sample of size n and $x_{1:n} \le x_{2:n} \le \dots \le x_{n:n}$ be an ordered sample. Then the r-th sample L-moment can be defined as

$$l_r = \binom{n}{r}^{-1} \sum_{1 \le i_1 < i_2 < \dots < i_r \le n} \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot x_{i_{r-j}:n}, \quad r = 1, 2, \dots, n. \tag{11}$$

Hence the first four sample L-moments have a form

$$l_1 = \frac{1}{n} \cdot \sum_i x_i, \tag{12}$$

$$l_2 = \frac{1}{2} \cdot \binom{n}{2}^{-1} \cdot \sum_{i > j}\sum (x_{i:n} - x_{j:n}), \tag{13}$$

$$l_3 = \frac{1}{3} \cdot \binom{n}{3}^{-1} \cdot \sum_{i > j > k}\sum\sum (x_{i:n} - 2x_{j:n} + x_{k:n}), \tag{14}$$

$$l_4 = \frac{1}{4} \cdot \binom{n}{4}^{-1} \cdot \sum_{i > j > k > l}\sum\sum\sum (x_{i:n} - 3x_{j:n} + 3x_{k:n} - x_{l:n}). \tag{15}$$

The natural estimate of the L-moment ratio (7) is the sample L-moment ratio

$$\tau_r = \frac{\lambda_r}{\lambda_2}, \quad r = 3, 4, \dots. \tag{16}$$

1.3 Parameter estimation

The random variable X has a three-parameter lognormal distribution with parameters μ, σ^2 and θ, where $-\infty < \mu < \infty$, $\sigma^2 > 0$, $-\infty < \theta < \infty$, if its probability density has the form

$$f(x;\,\mu,\,\sigma^2,\,\theta) = \frac{1}{\sigma\cdot(x-\theta)\cdot\sqrt{2\pi}}\cdot\exp\left[-\frac{[\ln(x-\theta)-\mu]^2}{2\sigma^2}\right], \qquad x > \theta, \tag{17}$$

$$= 0, \qquad \text{else.}$$

Let $\Phi^{-1}(\cdot)$ be distribution function of the standardized normal distribution. Parameter estimates obtained by the L-moment method for the case of a three-parameter lognormal distribution are obtained using the following equations[2] (“L” means L-moment estimation)

$$z = \sqrt{\frac{8}{3}}\cdot\Phi^{-1}\left(\frac{1+t_3}{2}\right), \tag{18}$$

$$\sigma^{\mathrm{L}} \approx 0{,}999\,281z - 0{,}006\,118z^3 + 0{,}000\,127\,z^5, \tag{19}$$

$$\mu^{\mathrm{L}} = \ln\left[\frac{l_2}{\operatorname{erf}\left(\frac{\sigma^{\mathrm{L}}}{2}\right)}\right] - \frac{\sigma^{2^{\mathrm{L}}}}{2}, \tag{20}$$

$$\theta^{\mathrm{L}} = l_1 - \exp\left(\mu^{\mathrm{L}} + \frac{\sigma^{2^{\mathrm{L}}}}{2}\right). \tag{21}$$

(Hosking, 1990)

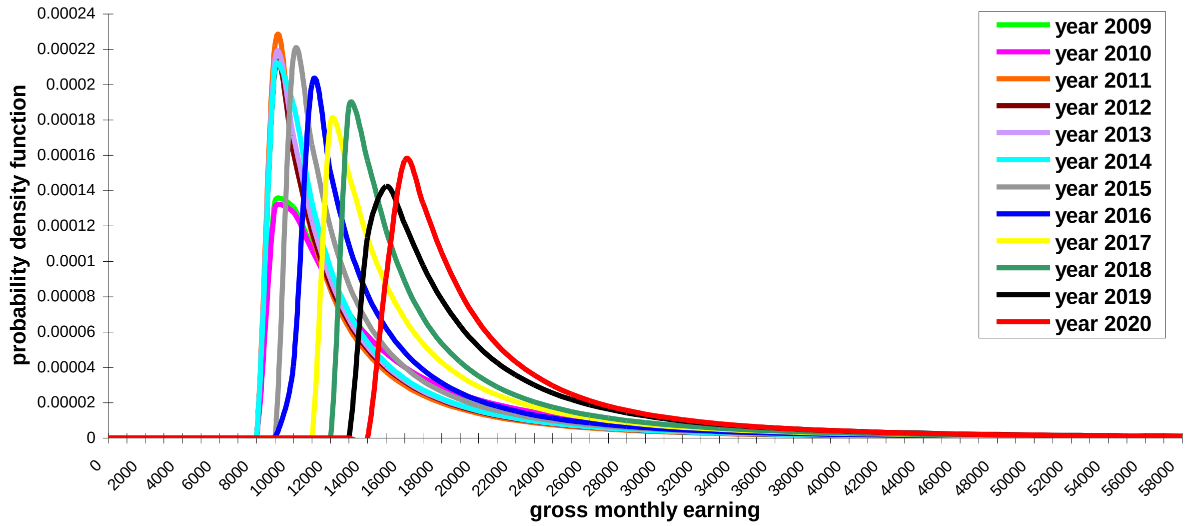
2 Results and Conclusion

Figures 1–2 offer a comparison of the development of the model distribution of earnings (for wages of employees in the private sphere and salaries of employees in the public sphere together) of the two sectors in which the Czech Republic has the lowest earnings, in the period 2009–2020. There are Accommodation and Food Service Activities and Sector of Administrative and Support Service Activities sectors. Figures 3–4 allow a comparison of development of model distribution of earnings between the two sectors over the years 2009–2020. Relative frequencies were calculated based on model distribution of earnings.

[2] Expression erf(z) is the so-called error function

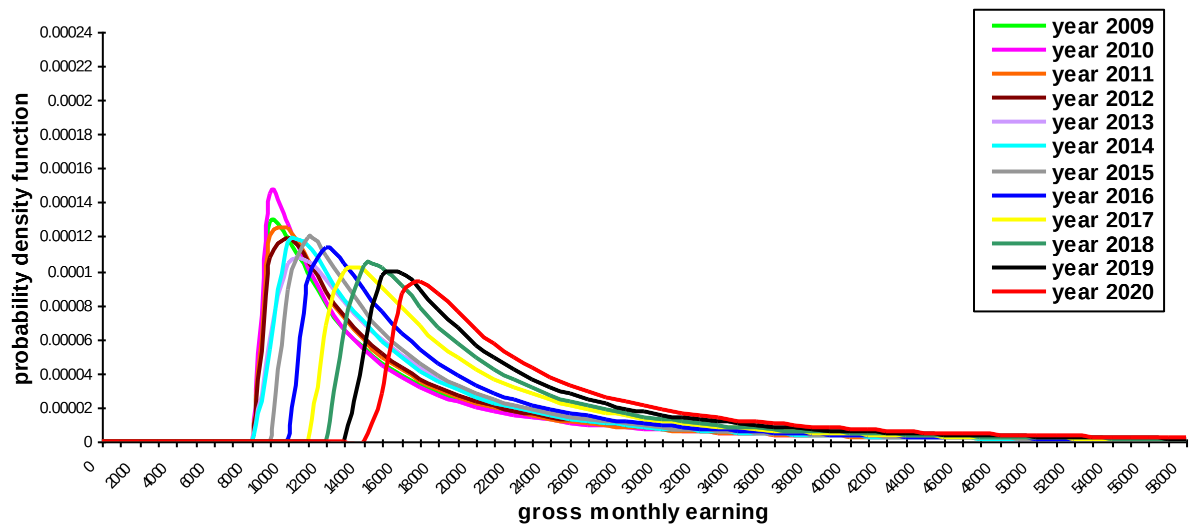
$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\cdot\int_0^z \exp(-t^2)\,\mathrm{d}t.$$

Fig. 1: **Development of model earning distribution of Accommodation and Food Service Activities sector**



Source: Own construction

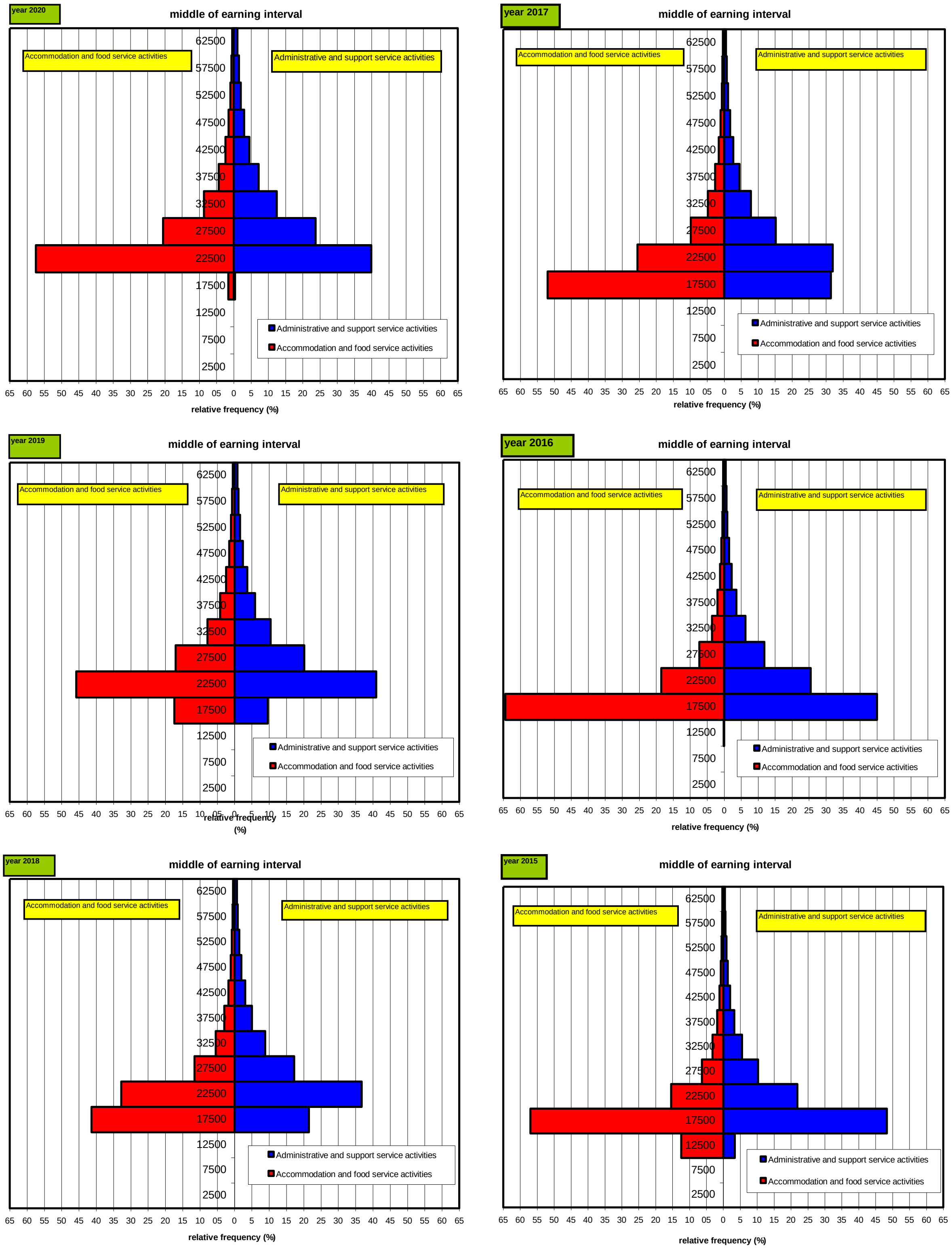
Fig. 2: Development of model earning distribution of Administrative and Support Service Activities sector



Source: Own construction

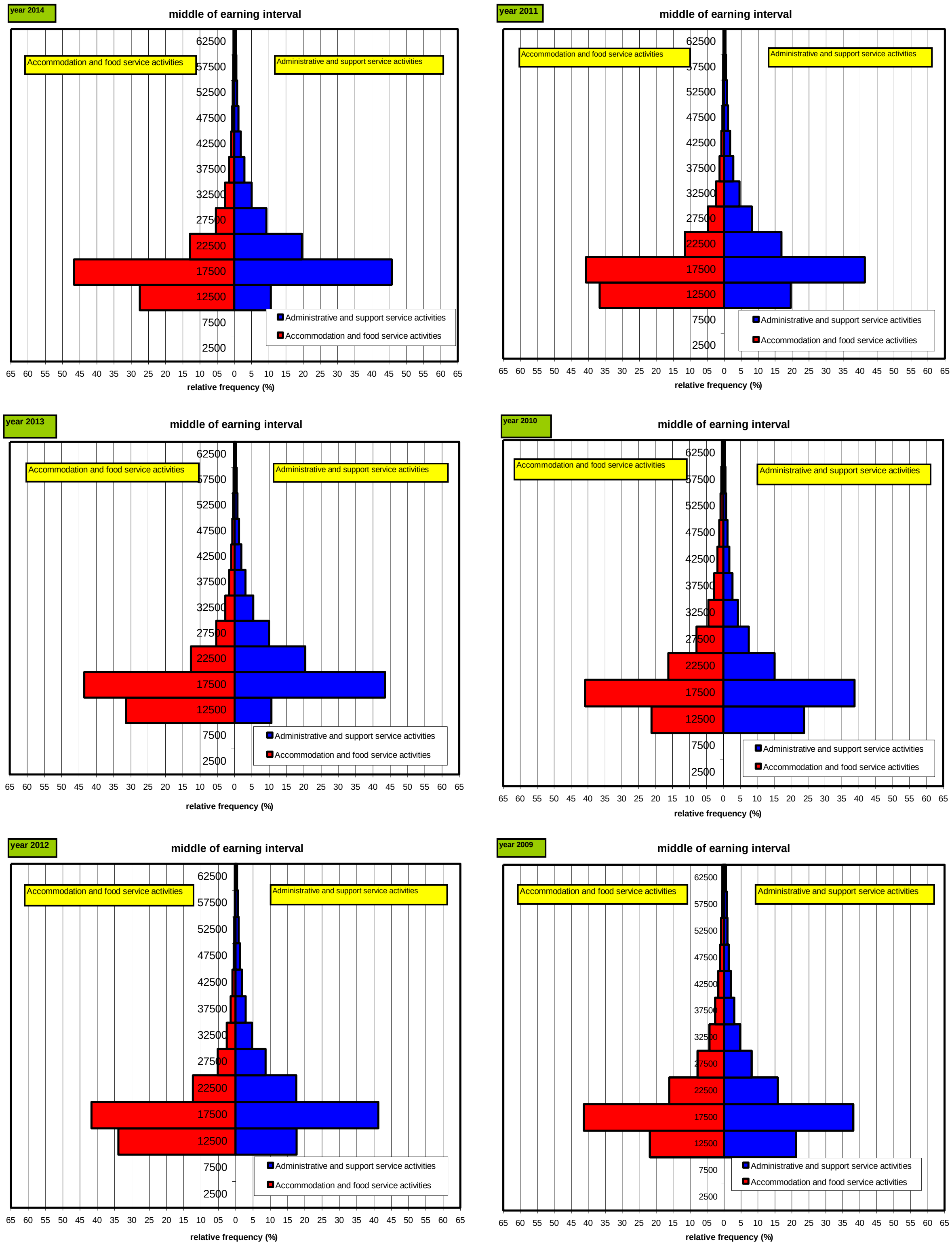
From Figures 1‒2, this is clear that the model distributions of earnings in the Accommodation and Food Services Activities sector are characterized by a lower level and variability compared to the model distributions of earnings in the Administrative and Support Service Activities sector. It is also clear that the model distributions of earnings in the Accommodation and Food Service Activities sector are more skewed, and they have more kurtosis than the model distributions of earnings in the Administrative and Support Service Activities sector.

Fig. 3: Comparison of the distribution of relative frequencies (%) of Sector of Accommosation and Food Service Activities and Sector of Administrative and Support Service Activities during 2015–2020



Source: Own construction

Fig. 4: Comparison of the distribution of relative frequencies (%) of Sector of Accommosation and Food Service Activities and Sector of Administrative and Support Service Activities during 2009–2014

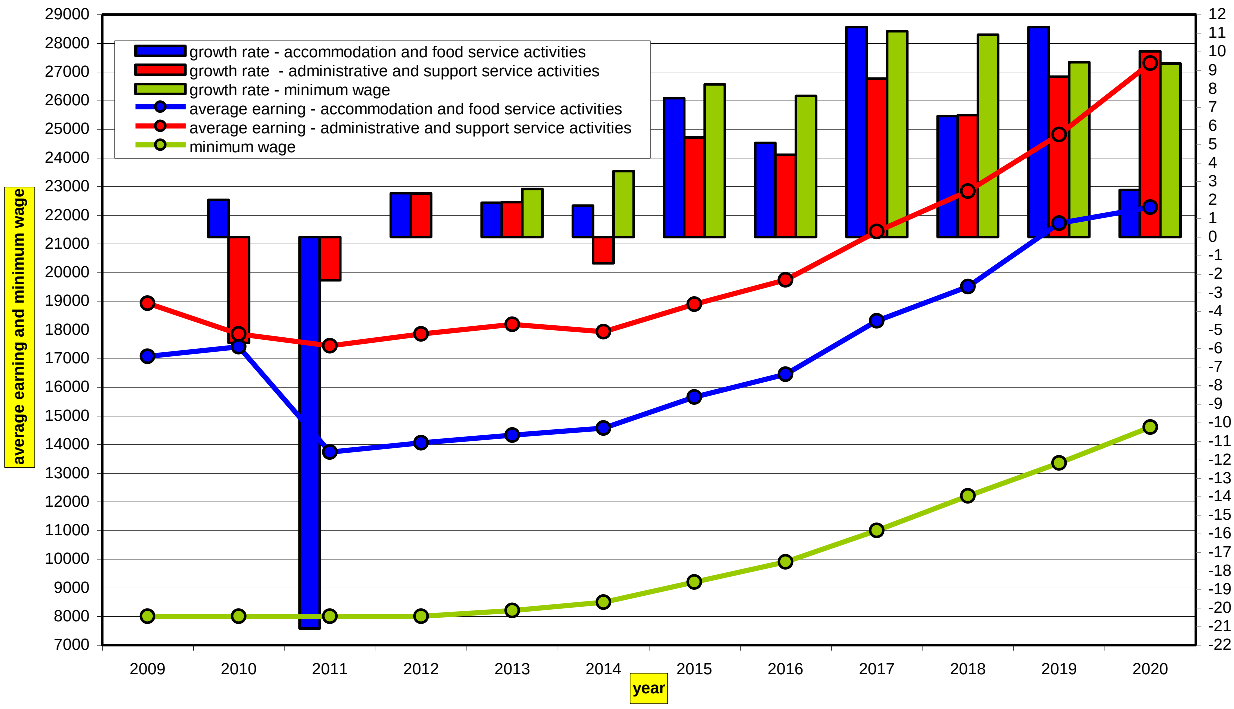


Source: Own construction

In Accommodation and Food Service Activities sector, the frequencies of lower earning intervals are higher, then in Administrative and Support Service Activities sector and vice versa, see Figures 3‒4.

Figure 5 shows the development of average earnings in each of the two sectors and the development of the minimum wage and the corresponding growth rates in the period 2009‒2020.

Fig. 5: Development of average earnings in Accomodation and Food Service Activities and Administrative and Support Service Activities sectors and minimum wage (CZK) and development of growth rate (%) of average earnings in Accomodation and Food Service Activities and Administrative and Support Service Activities sectors and minimum wage during period 2009–2020



Source: Own construction

The lower level of earnings in Accommodation and Food Services Activities sector compared to the Administrative and Support Services Activities sector can be seen in Figure 5, too.

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

Alahmadi, F. (2017). Regional Rainfall Frequency Analysis by L-Moments Approach for Madina Region, Saudi Arabia. *International Journal of Engineering Research and Development*, *13*(7), 39–48.

Eslamian, S. S. & Feizi, H. (2007). Maximum Monthly Rainfall Analysis Using L-Moments for an Arid Region in Isfahan Province, Iran. *Journal of Applied Meteorology and Climatology*, *46*(4), 494–503.

Guttman, N. B. (1994). On the Sensitivity of Sample L Moments to Sample Size. *Journal of Climate*, *7*(6), 1 026–1 029.

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)*, *52*(1), 105–124.

Hosking, J. R. M. (2006). On the Characterization of Distributions by Their L-Moments. *Journal of Statistical Planning and Inference*, *136*(1), 193–198.

Kumar, R. & Chatterjee, C. (2005). Regional Flood Frequency Analysis Using L-moments for North Brahmaputra Region of India. *Journal of Hydrologic Engineering*, *10*(1), 1–7.

Kyselý, J. & Picek, J. (2007). Regional Growth Curves and Improved Design value Estimates of Extreme Precipitation Events in the Czech Republic. *Climate research, 33*(3), 243–255.

Noto, L. V. & La Loggia, G. (2009). Use of L-Moments Approach for Regional Flood Frequency Analysis in Sicily, Italy. *Water Resources Management*, *23*(9), 2 207–2 229.

Pearson, C. P. (1991). New Zealand Regional Flood Frequency Analysis Using L-Moments. *Journal of Hydrology (New Zealand)*, *30*(2), 53–64.

Saf, B. (2009). Regional Flood Frerquency Analysis Using L-Moments for the West Mediterranean Region of Turkey. Water Resources Management, *23*(2), 531–551.

Contact

Diana Bílková
Prague University of Economics and Business
Faculty of Informatics and Statistics
Department of Statistics and Probability
W. Churchill Square 1938/4
130 67 Prague 3
Czechia
bilkova@vse.cz