ANALYSIS OF THE HIGHEST WAGES IN THE SLOVAK REPUBLIC

Viera Pacáková[a)], **František Foltán**[b)]

[a)]University of Pardubice, Faculty of Economics and Administration, Institute of Mathematics, [b)]University of Economics in Bratislava, Faculty of Economic Information, Department of Statistics

Abstract: *This article contains the results of statistical analysis of the highest wages in the Slovak Republic in 2009 using the methods of statistical inference. The information from the sample of the employees with wage higher than the 99th percentile is generalized to the population using the Pareto distribution of wages. There are also identified the factors that have a significant impact on the frequency and on the level of extreme wages. These factors are gender, level of education, industry and region of residence.*

Keywords: *Gros Wage, Pareto Distribution, Goodness-of-fit Tests, Quantiles.*

JEL Classification: *C46, C83, D14, D31, D33*

1. Introduction

Reliable information about wages in each country is very important of several economic and social reasons. Therefore, the data about incomes of persons or households are collected in number of sample surveys and obtained data are analyzed by various methods.

Most of such analysis is limited to a simple description of sample data. This gives unreliable information, especially on the highest incomes, which are rare in the sample.

Statistical inference methods provide a tool to generalize the information from the sample to the population that in our case is the set of all employees in the Slovak Republic (further referred to as SR) in 2009. Observed random variable is the gross monthly wage of the employees in SR.

In recent years in Slovakia and in the Czech Republic several authors have successfully dealt with the analysis and modeling of household income using the methods of statistical inference.

The results of such analysis are several interesting publications about the distribution of household's income, for example Bartošová, J. (2007), Sipková, Ľ. and Sipko, J. (2010), or Labudová, V. and Šoltés, E. (2008).

This article aims to obtain reliable information on wages of all employees in Slovakia in 2009, using methods of statistical inference. It focuses mainly on the highest wages, their distribution and the fundamental factors affecting the amount of those wages. These factors are gender, level of education, industry and region of residence. All analysis use the statistical software package Statgraphics Centurion XV.

2. Statement of a problem

The analysis of the highest gross monthly wages, particularly those beyond the 99th percentile of the sample uses the sample survey information system on labour costs, which has been realized by Trexima in the Slovak Republic since 1992. Therefore the basis of analyse is sampling data set of one percentage of employees in the number of 9 069 in SR, who earned a gross monthly wage of more than 3 488,76 EUR in 2009.

The most advanced generalization of information from the sample to population is to find the probability distribution of observed variable. In our case probability distribution provides full information on the wages of all employees of the SR with a gross monthly wage above 99th percentile of the sample. Knowledge of appropriate probabilistic model allows the calculation of the characteristics of the population, quantiles, probabilities of any intervals of wages and so on.

Pareto probability model with a lower bound a is suitable for modeling of gross monthly wages at the upper tail of distribution.

This distribution is named after economist W. Pareto (1848 – 1923), who used it to model welfare economics. We will use it as a probabilistic model of gross monthly wages above 3 488,76 EUR in Slovakia in 2009. This variable will be referred to as X_a.

When a random variable X_a has a Pareto distribution with parameters $a > 0$ and $b > 0$, its distribution function is given by

$$F_a(x) = 1 - \left(\frac{a}{x}\right)^b, \qquad x \geq a \tag{1}$$

where $a > 0$ is threshold parameter and $b > 0$ is shape parameter [5, p. 20].

The mean value is expressed as follows:

$$E(X_a) = \frac{a \cdot b}{b - 1}, \qquad b \rangle 1 \tag{2}$$

and the variance

$$D(X_a) = \frac{a^2 \cdot b}{(b-1)^2 \cdot (b-2)}, \qquad b \rangle 2 \tag{3}$$

3. Problem solving

3.1. Sample statistics

In the sample, which consists of 9 069 employees with the highest 1 % gross monthly wage (over 3 488,76 EUR), average gross monthly wage is 6 089,28 EUR, half of the employees earns more than the median level 4 763,50 EUR, gross monthly wage of a quarter of employees exceeded 6 497,96 EUR and the maximum wage is 5 9648,30 EUR. Existence of outliers is confirmed by the high value of coefficient of variation, that is 72,35 %.

A visual form of the layout values of gross monthly wages above 99th percentile in the sample provides a box plot in Fig. 1.
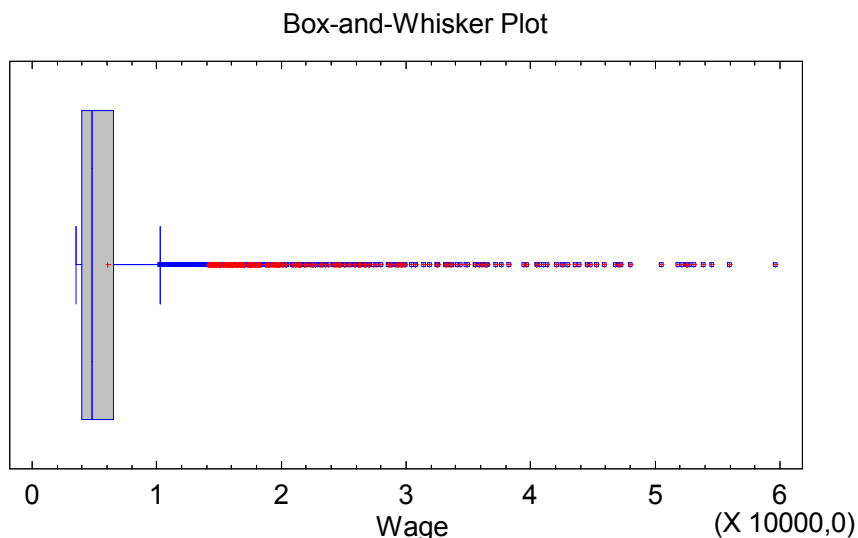
Box-and-Whisker Plot

Fig. 1: Box plot of gross monthly wages higher than 99th percentile

3.2. Pareto distribution of the highest gross monthly wages

Distribution fitting procedure in the statistical software package Statgraphics Centurion XV allows by the sample data, exceeding a certain value *a*, to estimate the parameters *a*, *b* of Pareto distribution by maximum likelihood method (Pacáková, 2004: 67-68) and to verify with help of the seven goodness of fit tests, whether sample data may originate from such distribution. We have used Kolmogorov-Smirnov test (Pacáková 2004: 82-84). This test has confirmed good fit with the Pareto distribution with parameters *a* = 3 488,76 and *b* = 2,28473 at significance level 0,05 *(p*-value 0,08686> 0,05), as follows from the distribution fitting procedure output at Fig. 2.

9069 values ranging from 3488,76 to 59648,30

Fitted Distributions

Goodness-of-Fit Tests for wage

Kolmogorov-Smirnov Test

Pareto (2-Parameter)
shape = 2,28473
lower threshold = 3488,76

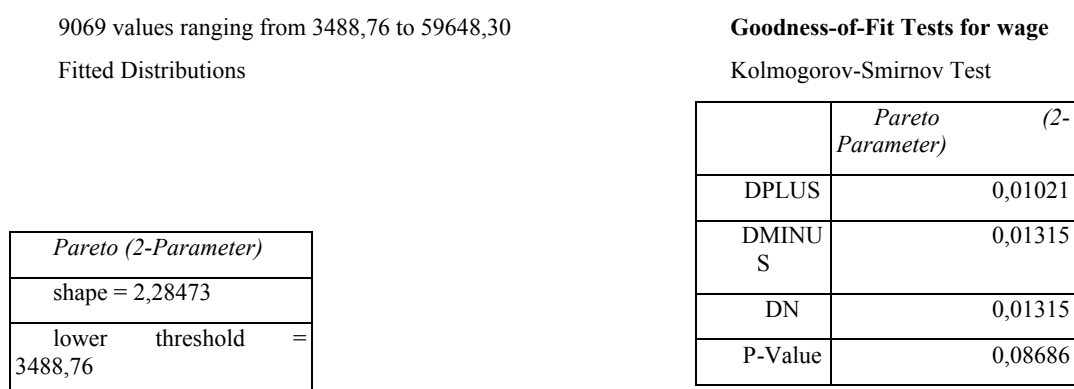|	Pareto (2-Parameter)
DPLUS	0,01021
DMINUS	0,01315
DN	0,01315
P-Value	0,08686

Fig. 2: Result of goodness of fit test with Pareto distribution

From Fig. 3, that shows a histogram of sample data together with Pareto density function is evident, that the sample data may originate from the fitted distribution.
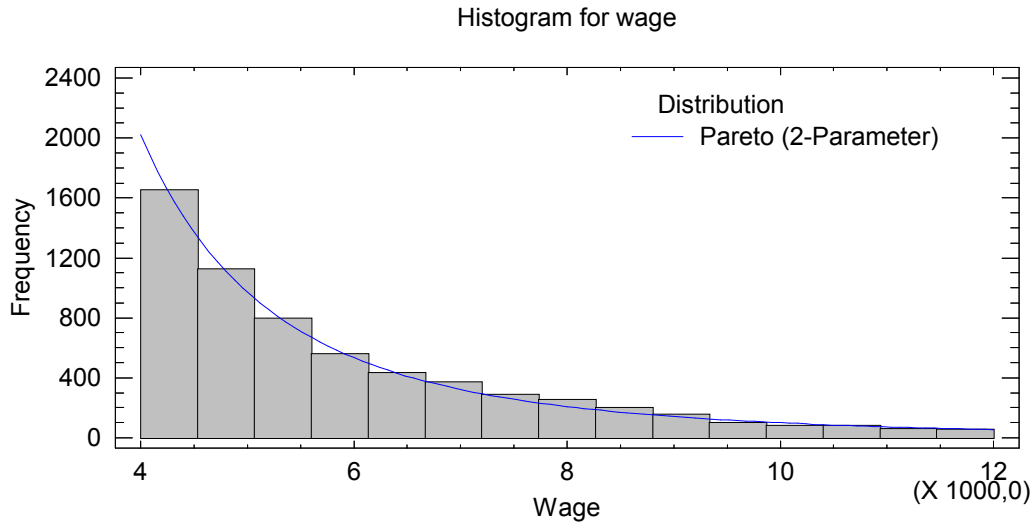
174

Fig. 3: Graphical verification of good fit with Pareto distribution

Statgraphics Centurion XV allows us to find any percentile of gross monthly wages distribution, as we can see on the output in Tab. 1. These quantiles give us valuable information, for example that 5 % of employees in SR in 2009 with a gross monthly wage above 3 488,76 EUR received a wage higher than 12 945,40 EUR. According to equation (2) we can obtain the basic characteristics of the Pareto distribution: $E(X_a) = 6\ 204,32$ and $D(X_a) = 59\ 172\ 526,63$.

Tab.1: Selected quantiles of gross monthly wages above 99th percentile in SR in 2009

Lower Tail Area (<=)	Pareto (2-Parameter)
0,10	3653,41
0,50	4725,28
0,75	6400,08
0,90	9557,81
0,95	12945,40

3.2 Factors affecting the amount of the wages above 99th percentile

The method of the analysis of variance allows verifying significance of the factors *gender*, *education*, *industry* and *residence region of the employee* on the amount of the wages above 99th percentile in Slovakia in 2009.

Tab. 2: Sample statistics of the highest wages by gender

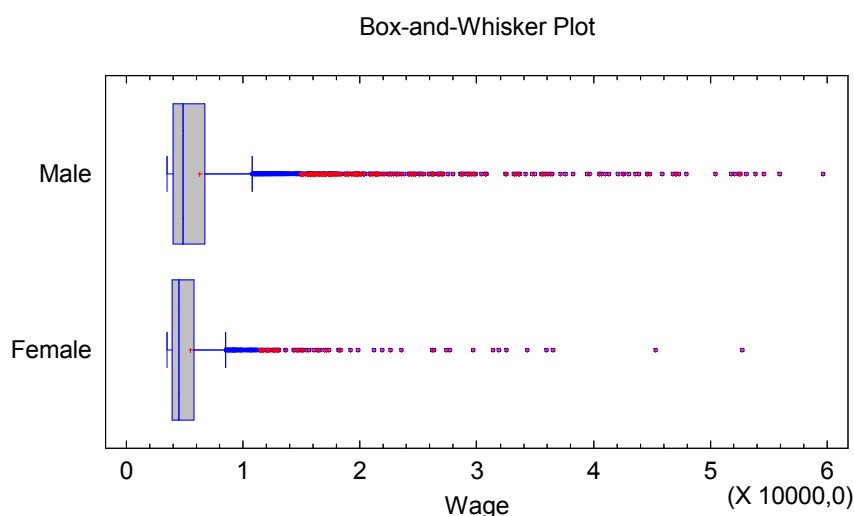| Gender	Count	Average	Median	Coefficient of variation	Minimum	Maximum	Upper Quartile	5/6 Sextile
1-male	7185	6253,61	4863,15	74,04 %	3488,76	59648,30	6729,33	7916,67
2-female	1884	5462,57	4526,21	61,20 %	3489,31	52739,10	5747,64	6669,29
Total	9069	6089,28	4763,50	72,35 %	3488,76	59648,30	6497,96	7678,39

175

Box-and-Whisker Plot

Fig. 4: Box plots of the highest monthly wages by gender

The sample statistics in Tab. 2 (1-male and 2-female) shows lower percentage of female (only 20,77 %) and also lower female wage level in a set of 1 % of employees with the highest wages in Slovakia in 2009.

Tab. 3: Sample statistics of the highest wages by education

Education	Count	Percent	Average	Median	Coefficient of variation	Minimum	Maximum	Upper Quartile	5/6 Sextile
0	994	10,96	6688,77	4930,87	90,88 %	3492,12	55941,40	7011,10	8498,14
1	22	0,24	5086,74	4550,82	25,53 %	3625,97	8262,28	6143,80	6689,12
2	80	0,88	5644,04	4310,28	54,57 %	3489,50	18060,00	6021,78	7613,49
3	46	0,51	4940,27	4337,02	37,95 %	3505,69	15059,60	5311,03	6032,58
4	253	2,79	5321,23	4401,98	63,99 %	3489,62	34952,60	5726,46	6486,21
5	219	2,41	5877,48	4757,20	69,70 %	3492,70	47079,10	6195,42	7840,67
6	529	5,83	5689,18	4613,92	68,44 %	3489,03	45881,10	6293,19	7166,24
7	38	0,42	6286,36	4932,55	86,40 %	3512,28	35929,70	5749,70	6436,39
8	220	2,43	6153,66	4806,05	63,68 %	3489,31	35764,10	6387,78	7739,31
9	6530	72,00	6065,09	4772,69	68,57 %	3488,76	59648,30	6514,17	7673,40
10	138	1,52	6837,72	4711,96	93,76 %	3501,89	47021,70	7322,36	8457,01
Total	9069	100,00	6089,28	4763,50	72,35 %	3488,76	59648,30	6497,96	7678,39

Education factor has 11 observed levels: 0-not reported, 1-primary, 2-apprenticeship, 3-secondary (without exam), 4-apprenticeship with leaving exam, 5-full secondary general, 6-full secondary vocational, 7-higher, 8-university 1st degree (bachelor), 9-university 2nd degree, 10-university 3rd degree (research qualification).

Tab. 3 contains sample statistics and the percentage of employees on the different educational level.

Box-and-Whisker Plot

Fig. 5: Box plots of the highest monthly wages by education

The largest group (72 %) consists of university educated employees (9-university 2nd degree) in the set of 1 % employees with the highest gross monthly wages in Slovakia in 2009. The highest educated employees (research qualification) have surprisingly low representation, only 1,52 %, in this set of employees.

Tab. 4: Sample statistics of the highest wages by selected industries

SK NACE	Count	Percent	Average	Median	Coefficient of variation	Minimum	Maximum	Upper Quartile	5/6 Sextile
C	2190	24,15	6004,92	4796,78	63,28 %	3488,80	50447,70	6535,27	7587,77
D	468	5,16	5797,32	4897,14	51,60 %	3490,40	30787,00	6570,94	7557,77
F	334	3,68	6779,29	4985,54	85,76 %	3493,40	59648,30	6934,27	8701,92
G	1117	12,32	6354,32	4868,30	77,81 %	3489,04	47280,60	6644,72	7913,13
H	525	5,79	5856,04	4886,15	59,12 %	3494,00	47968,10	6790,31	7464,81
J	1666	18,37	5872,25	4585,82	74,90 %	3489,15	54546,20	6072,22	7257,35
K	1339	14,76	6100,91	4742,24	66,04 %	3488,76	47079,10	6526,89	8098,74
M	677	7,46	6930,83	5098,46	96,60 %	3489,59	55941,40	7011,94	8202,55
N	110	1,21	7486,50	5162,31	87,49 %	3503,24	41015,00	8443,92	9818,66
O	127	1,40	5137,31	4145,51	48,60 %	3506,82	15725,00	5035,82	6245,90
Q	185	2,04	5138,24	4177,94	50,72 %	3496,15	21960,70	5111,31	5902,50
Total	9069	100,00	6089,28	4763,50	72,35 %	3488,76	59648,30	6497,96	7678,39

Industry factor in the sample has 21 modifications. Employees with high salaries are not represented in all sectors. Their representation in some sectors is negligible compared to the other.

Codes of *selected industries* are as follows: C-manufacturing, D-electricity, gas, steam and air conditioning supply, F-construction, G-wholesale and retail trade, H-transportation and storage, J-information and communication, K-financial and insurance activities, M-professional, scientific and technical activities, N-administrative and support service activities, O-public administration and defence, Q-human health and social work activities. The greatest proportion of employees with the highest monthly wages, nearly one quarter, is in industry C-manufacturing. High proportion of employees with salaries above 99 percentile is also in industries J-information and communication, K-financial and insurance activities and G-wholesale and retail trade. More detailed information is provided by Tab. 4 and Fig. 6.

Fig. 6: Box plots of the highest monthly wages by industry

Significant differences in the number of employees with gross monthly wages beyond the 99th percentile as well as in the amount of their wages causing also factor *the region of residence*. Codes of regions are as follows: 1-Bratislava region, 2-Trnava region, 3-Trenčín region, 4-Nitra region, 5-Žilina region, 6-Banská Bystrica region, 7-Prešov region, 8-Košice region.

Tab. 5: Sample characteristics of the highest wages by region

Region	Count	Percent	Average	Median	Coefficient of variation	Minimum	Maximum	Upper Quartile	5/6 Sextile
1	5679	62,62	6209,62	4805,23	74,30 %	3488,76	55941,40	6655,72	7887,67
2	523	5,77	6450,69	4771,91	83,46 %	3489,03	44569,70	6700,86	7691,82
3	534	5,89	5761,52	4675,73	70,23 %	3491,03	59648,30	6320,17	7055,91
4	453	5,00	5780,06	4643,21	66,11 %	3490,71	41331,50	6090,74	7002,83
5	631	6,96	6037,64	4869,17	60,85 %	3491,32	30782,20	6569,30	7725,63
6	386	4,26	5504,97	4541,74	55,99 %	3489,15	41015,00	5791,46	6958,45
7	256	2,82	5924,64	4756,76	62,70 %	3495,94	30416,30	6184,32	6998,15
8	607	6,69	5665,74	4691,77	66,06 %	3489,82	44776,10	6015,35	7017,68
Total	9069	100,00	6089,28	4763,50	72,35 %	3488,76	59648,30	6497,96	7678,39

The highest proportion of employees with the highest gross monthly wage was in 1-Bratislava region (62,62 %). In other regions is proportional representation comparable, the lowest is in 7-Prešov Region. More detailed information is provided by Tab. 5 and Fig. 7.

Fig .7: Box plots of the highest monthly wages by region

4. Conclusion

Despite the worldwide crisis which has been reflected also in the economic results and living conditions in the Slovak Republic there is a group of employees with very high level of gross monthly wages. More information about this group of employees we obtained using statistical methods.

We have confirmed the significance of the factors gender, education, industry and region of residence on the highest wages and also the fact that extremely gross

monthly wages are distributed by Pareto distribution. The highest wages receive mainly male employees with 2nd degree university education, working in the industry C-Manufacturing and living in the Bratislava region.

References

[1]   BARTOŠOVÁ, J. Pravděpodobnostní model rozdělení příjmů v České republice. *In Acta Oeconomica Pragensia*, 2007,15(1), s. 7-12.

[2]   PACÁKOVÁ, V. *Aplikovaná poistná štatistika.* Bratislava: IURA EDITION, 2004. ISBN 80-8078-004-8.

[3]   SIPKO, J. - SIPKOVÁ, Ľ. Aktuálny vývoj príjmovej nerovnosti na Slovensku. *In Forum statisticum Slovacum*, Bratislava, 2010. ISSN 1336-7420, 2010, roč. 6, č. 5, s. 218-223.

[4]   ŠOLTÉS, E. - LABUDOVÁ, V. Výber relevantných faktorov a štatistické posúdenie ich vplyvu na disponibilný príjem domácností SR na základe zisťovania SILC 2006. *In Ekonomika a informatika*, 2008, 6(1/2008), s. 88-99.

[5]   STATGRAPHICS Centurion XV, On-line manual by StatPoint, Inc. Probability Distributions, 2005, p. 20.

Contact Address

prof. RNDr. Viera Pacáková, Ph.D.
University of Pardubice, Faculty of Economics and Administration, Institute of Mathematics
Studentská 84, 532 10 Pardubice, Czech Republic
E-mail: vpacakova@gmail.com
Phone number: +420 466 036 019

Ing. František Foltán
University of Economics in Bratislava, Faculty of Economic Information, Department of Statistics
Dolnozemská cesta 1/b, 852 35 Bratislava, Slovak Republic
E-mail: frantisek.foltan@euba.sk
Phone number: +421 2 672 95 725