Work Intensity of Households: Multinomial Logit Analysis and Correspondence Analysis for Slovak Republic

Erik Šoltés [1] | *University of Economics in Bratislava, Bratislava, Slovakia*
Mária Vojtková [2] | *University of Economics in Bratislava, Bratislava, Slovakia*
Tatiana Šoltésová [3] | *University of Economics in Bratislava, Bratislava, Slovakia*

Abstract

Exclusion from the labour market is a serious social problem that is also addressed by the Europe 2020 strategy. While in the past the attention of statisticians and sociologists in the fight against poverty and social exclusion has concentrated mainly on income poverty and material deprivation, in recent times many studies and analyses are much more focused on work intensity as well. Households that use their work potential to less than 20%, have a very low work intensity, and members of such households are included into the population of people who are at risk of poverty or social exclusion. Moreover, the low use of labour potential of households significantly increases the risk of income poverty and the threat of material deprivation. This article provides an analysis of work intensity levels of Slovak households depending on the factors that are monitored by the EU-SILC 2015. The impact of relevant factors is quantified by correspondence analysis and by multinomial logistic regression model.

Keywords	***JEL code***
Work intensity, poverty and social exclusion, EU-SILC – European Union statistics on income and living condition, correspondence analysis, multinomial logistic regression	*I31, J21, C31*

INTRODUCTION

One of the fundamental objectives of the Europe 2020 strategy is to reduce poverty and social exclusion in the EU and its member states. In order to monitor the achievement of this goal, an aggregated indicator measuring the risk of poverty or social exclusion (AROPE) was created. The key objective in the social field

[1] University of Economics in Bratislava, Faculty of Economic Informatics, Department of Statistics, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia. E-mail: erik.soltes@euba.sk, phone: (+421)267295717.
[2] University of Economics in Bratislava, Faculty of Economic Informatics, Department of Statistics, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia. E-mail: maria.vojtkova@euba.sk, phone: (+421)267295724.
[3] University of Economics in Bratislava, Faculty of Economic Informatics, Department of Mathematics and Actuarial Science, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia. E-mail: tatiana.soltesova@euba.sk, phone: (+421)267295835.

is to pull 20 million people out of the risk of poverty or social exclusion by 2020 compared to the year 2008, while people are considered to be at risk of poverty or social exclusion if they are at risk of income poverty and / or are materially deprived and / or living in households with very low work intensity. Therefore, the methodology for measuring poverty and social exclusion used in the Europe 2020 strategy, is based on a 3-dimensional concept with the following dimensions: income poverty, material deprivation and exclusion from the labour market.

The article focuses on the third dimension, which is mapped through the work intensity. Work intensity reflects the extent to which the working potential is used in the household, in other words, how much of the theoretically available work time (set in the legislation of the country) the household members in productive age actually work. The household work intensity can have value from 0 (or 0% – no one works) up to 1 (or 100% – all members work). On this basis, the indicator "Share of people living in very low work intensity households" was created, it refers to people aged 0–59 living in households where adults aged 18–59 work less than 20% of their overall work potential. The Europe 2020 strategy mainly tracks very low work intensity and, therefore, the professional and scientific community monitors and analyses the intensity of work of people and households, mainly through measuring very low work intensity. Despite the fact that several scientific publications point out the weaknesses of this indicator, e.g. Ward and Ozdemir (2013) and the delayed disclosure of its estimated values in connection with the complicated nature of EU-SILC micro-data collection and processing (Rastrigina et al., 2015), the very low work intensity rate and the measurement methodology[4] of personal or household work intensity provide important internationally comparable information about exclusion from the labour market. Several authors include unemployment rate and very low work intensity rate among relevant indicators of economic welfare in the countries. For example, Monte et al. (2017) applied very low work intensity rate for monitoring the development of welfare in Europe between 2007 and 2014, while using the cluster analysis. On the basis of the analyses, the Slovak Republic entered together with the Czech Republic, Hungary and Slovenia in a joined cluster in all the years under review, and in 2014 the cluster of these countries reached a relatively low rate of very low work intensity.

This article does not only focus on very low work intensity but our goal is to provide a comprehensive analysis of work intensity of Slovak households, not only in terms of the incidence and risk of very low work intensity, but with regard to all degrees of work intensity.[5] To achieve this goal, we use correspondence analysis and logistic regression in this article. Logistic regression is a popular statistical tool in work intensity analysis and was also used by Mysíková et al. (2015) who demonstrated that work intensity in the SR and the Czech Republic during the period 2006–2013 significantly influenced the risk of income poverty. Through the logistical regression Kis and Gábos (2016) confirmed the significant impact of work intensity on consistent poverty in the EU, Hick and Lanau (2017) quantified the impact of selected factors on in-work poverty and examined the impact of risk factors on very low work intensity in Ireland. We would like to note that Ireland has a long-term history of the highest incidence of very low work intensity in the EU (see: Šoltés and Šoltésová, 2016; Monte et al., 2017). Considering this fact, it is not surprising that a higher number of scientific publications focus on this problem in this country, for example, (Redmond, 2016), (Whelan and Maître, 2014) or (Logue and Callan, 2016). Several of the above-mentioned scientific publications show that many studies focus on analysing mutual relationship between the different dimensions of the concept of measuring poverty and social exclusion.

The use of labour potential in many EU countries significantly affects the occurrence and the risk of poverty as well as material deprivation. Ayllón and Gábos (2015) confirmed the relationship between severe material deprivation and low work intensity in Central and Eastern Europe, whereas in other parts of Europe

4 See: <http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_EU-SILC)_methodology_%E2%80%93_concepts_and_contents#Work_intensity_.28WI.29>.

5 See: <http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_definition_of_dimensions>.

they did not demonstrate such dependence. The strong positive relationship between very low work intensity and poverty has been quantified by the authors in all analysed countries. Guagnano et al. (2013) has again revealed through correspondence analysis that work intensity is one of the major socio-economic factors influencing the perception of subjective poverty in Europe.

1 METHODS OF ANALYSIS

The article focuses on the degree of work intensity of Slovak households, which is recorded in the form of a multinomial categorical variable for individual statistical units (households). In regard to the character of the target variable and the goal of this article stated in the introduction, the results presented in this article are achieved by means of correspondence analysis and analysis of the multinomial logistic model. This part of the article provides a brief description of the methodology of these sophisticated mathematical and statistical methods.

1.1 Correspondence analysis

Correspondence analysis is a method that is based on the analysis of the structure of mutual dependencies of two or more variables. Because it focuses on examining the dependence of predominantly nominal or ordinal variables, in the case of a continuous variable, it is necessary to categorize its values. It solves this problem in a similar way as factor analysis or the principal component method, while hidden or latent variables can be represented as axes of the reduced coordinate system (correspondence maps), in which the individual categories of variables will eventually be displayed. This is a method that in its essence belongs to exploration methods, and can be a good instruction for deciding which categories of variable should be merged and which can be kept separate. It is mainly used in marketing, but its interesting applications are also found in other areas.

In the case of a simple correspondence analysis (Řezanková, 2007; Meloun and Militký, 2012), we deal with a two-dimensional contingency table. From the values of this table (n_{ij}) we can deduce the correspondence matrix P with the elements p_{ij}

$$p_{ij} = \frac{n_{ij}}{n}, \tag{1}$$

where $i = 1, 2, ..., r$ and $j = 1, 2, ...$ s.

Row marginal relative frequencies p_{i+} are called row loads (r_i), with their line percentages being referred to as row profiles. Similarly, column marginal relative frequencies p_{+j} are called column loads (c_j), with their column percentages being referred to as column profiles. The whole correspondence matrix can be schematically expressed as follows:

$$\begin{bmatrix} \mathbf{P} & \mathbf{r} \\ \mathbf{c}^{\mathrm{T}} & 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} & r_1 \\ p_{21} & p_{22} & \cdots & p_{2s} & r_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rs} & r_r \\ c_1 & c_2 & \cdots & c_s & 1 \end{bmatrix}, \tag{2}$$

where $\mathbf{c}$ is the s-elements vector of the column loads and $\mathbf{r}$ is the r-elements vector of the row loads.

Each row (column) of the correspondence matrix can be represented as a point in a s-dimensional (r-dimensional) space with coordinates corresponding to the values of the respective profiles. We can then calculate the distances between individual points, while the most commonly used is the chi-squared distance between the i-th and the i'–th line produced by the relation:

$$\chi^2 = \sqrt{\sum_{j=1}^{s} \frac{\left(r_{ij} - r_{i'j}\right)^2}{c_j}}\ , \tag{3}$$

where r_{ij} are the elements of the row profiles matrix $\mathbf{R}$ and c_j weights correspond to the elements of the column load vector $\mathbf{c}^{\mathrm{T}}$. Similarly, we proceed in computing the differences (dissimilarities) between column categories.

The goal of the method **i**s to reduce the multidimensional space of vectors of row and column profiles, while maximally preserving the information contained in the original data. Usually, a two-dimensional space is used, i.e. plane. The point that lies directly in the plane and is closest to the corresponding point in space is called projection. The solution comes from a matrix $\mathbf{Z}$ of standardized residuals with elements:

$$z_{ij} = \frac{p_{ij} - p_{i+}p_{+j}}{\sqrt{p_{i+}p_{+j}}} \tag{4}$$

and its singular decomposition according to relationship:

$$\mathbf{Z} = \mathbf{U} \cdot \mathbf{\Gamma} \cdot \mathbf{V}^{\mathrm{T}}, \tag{5}$$

where $\mathbf{\Gamma}$ is the diagonal matrix and where the relationship $\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$ applies.

Prior to the estimation of the co-ordinates of each category, the choice of the normalization method should be made, i.e. the way to show points in the correspondence map. The so-called symmetric normalization is most commonly used in which we are interested in the mutual comparison of both row and column categories. In interpreting the results, the points are considered closer when there is a higher similarity between the corresponding categories.

1.2 Multinomial Logit Analysis

The logistic regression model is a special case of the general linear model (see: Ramon et al., 2010) and serves to model the categorical dependent variable depending on the explanatory variables of the continuous or categorical type. The binary logistic regression uses the logarithmic transformation of the odds of probability p for the desired event to occur $(y_i = 1$; the event that is being examined) to the probability 1–p of occurrence of the undesired event ($y_i = 0$). The natural logarithm of the odds is called logit and, unlike probability p, acquires any real values and can be modelled by a linear regression model (Stankovičová and Vojtková, 2007):

$$\operatorname{logit}(p_i) = \ln \frac{p_i}{1 - p_i} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}\ , \tag{6}$$

where p_i is the probability, so that $y_i = 1$ ($i = 1, 2, \ldots, n$), then $\beta_0, \beta_1, \ldots, \beta_k$ are the parameters of the logit model and $x_{i1}, x_{i2}, \ldots, x_{ik}$ are the values of the explanatory variables $X_1, X_2, \ldots, X_k$ which are observed for the i-th statistical unit. To obtain maximum likelihood estimators of parameters of the logistic regression model the Newton-Raphson algorithm is generally used (see: Allison, 2012).

After estimating the logistic model, it is important to verify its statistical significance and also verify whether the influence of the individual explanatory variables on probability p is significant. The significance of a logistic regression model is performed by a zero hypothesis test $\boldsymbol{\beta}^{\mathrm{T}} = (\beta_1, \beta_2, \ldots, \beta_k) = \mathbf{0}^{\mathrm{T}}$ against an alternative hypothesis – at least one regression coefficient is non-zero, while three different test chi-square statistics are mostly used (Likelihood ratio, Score statistics, Wald statistics). Allison (2012) discusses the differences between these statistical methods and at the same time notes that in large samples, there is no reason to prefer any of these statistics and they will generally be quite close in value.

In order to validate the significance of the explanatory variable influence, a Wald test is used. It tests the zero hypothesis showing that the respective explanatory variable does not affect the probability of occurrence of the explored event. To verify the hypothesis, Wald statistic:

$$Wald = \hat{\boldsymbol{\beta}}^{\mathrm{T}} \cdot \mathbf{S}_{\mathbf{b}}^{-1} \cdot \hat{\boldsymbol{\beta}}, \tag{7}$$

is used, where $\hat{\boldsymbol{\beta}}$ is the vector of regression coefficients estimates that stand at dummy variables for the respective factor (categorical explanatory variable) and $\mathbf{S_b}$ is the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$. Wald statistic has asymptotically χ^2 distribution with degrees of freedom equal to the number of parameters estimated for a given effect. A special case of the above test is the Wald test, which verifies the statistical significance of one regression coefficient. In this case Wald statistics is asymptotically distributed as χ^2 with 1 degree of freedom. The test statistic has an equation:

$$Wald = \frac{\hat{\beta}_i}{s_{\hat{\beta}_i}}, \tag{8}$$

where $s_{\hat{\beta}_i}$ is an estimated standard error of the i-th estimated coefficient.

The quality of the logistic model can be evaluated by different measures. Among criteria that measure a relative quality of statistical models belong AIC – Akaike Information Criterion and SC – Schwarz-Criterion, which are based on the logarithmic transformation of the likelihood function, i.e. $-2\ln L$.

Binary logistic regression is used, if the explanatory variable is binomial. If the dependent variable has more than 2 categories (generally these are s categories), we can use a multinomial logit model that is created by $(s - 1)$ logit functions:

$$\begin{aligned}
&\ln\left[\frac{P(Y=1|\mathbf{x})}{P(Y=0|\mathbf{x})}\right] = \beta_{10} + \beta_{11}x_{i1} + \beta_{12}x_{i2} + \cdots + \beta_{1k}x_{ik} \\
&\ln\left[\frac{P(Y=2|\mathbf{x})}{P(Y=0|\mathbf{x})}\right] = \beta_{20} + \beta_{21}x_{i1} + \beta_{22}x_{i2} + \cdots + \beta_{2k}x_{ik} \\
&\vdots \\
&\ln\left[\frac{P(Y=(s-1)|\mathbf{x})}{P(Y=0|\mathbf{x})}\right] = \beta_{(s-1)0} + \beta_{(s-1)1}x_{i1} + \beta_{(s-1)2}x_{i2} + \cdots + \beta_{(s-1)k}x_{ik}
\end{aligned} \tag{9}$$

The effect of the explanatory variable X_j on the dependent variable Y is quantified in logistic regression by the odds ratio (OR – odds ratio) estimated by the formula:

$$OR_j = e^{\hat{\beta}_j}. \tag{10}$$

The odds ratio in binary logistic regression expresses how the odds will change: $Y = 1$ compared to the odds that $Y = 0$, in unit growth of the explanatory variable in conditions ceteris paribus. If the explanatory variable is a dummy variable, the odds ratio compares the odds of occurrence of an event at two different levels of the predictor. In the case of multinomial logistic regression, the odds ratio interpretation is analogous to that of binomial logistic regression, we only have to consider which logit equation from Formula (9) we should take into account, and, therefore, which pair of categories of the multinomial explanatory variable we should compare (most often it is l vs. 0, where $l = 1, 2, \ldots, (s - 1)$).

2 DATABASE

The analyses presented in this article are based on the EU-SILC 2015 database provided by the Statistical Office of the Slovak Republic which covers the 2014 reference period. According to the methodology used by Eurostat to monitor labour market exclusion as one of the dimensions of poverty and social exclusion, work intensity is divided into 5 categories (Table 1). For the purpose of analysing the work intensity of Slovak households, we created a categorical variable *WI* (Work Intensity) whose variations (0 to 4) express the severity of the reduced use of households' work potential, as shown in Table 1.

Table 1 Levels of household work intensity

Level of work intensity	Value ranges of work intensity index	Category designation	Abbreviation
Very low	⟨0; 0.2)	4	VLWI
Low	⟨0.2; 0.45)	3	LWI
Medium	⟨0.45; 0.55⟩	2	MWI
High	(0.55; 0.85⟩	1	HWI
Very high	(0.85; 1⟩	0	VHWI

Source: Eurostat, own processing

According to the EU-SILC 2015 surveys in the reference year 2014, households with very high labour intensity dominated in Slovakia. There were 56.2% of such households (households that use more than 85% of their work potential) in the selected sample and 93% of them used their work potential up to 100%. 15.4% of households in this sample showed a medium degree of work intensity. The lowest number of households (3.9%) had a low level of work intensity (we included households that use their work potential to at least 20% but not more than 45% into this group). Almost every thirteenth household had to face a very low work intensity, respectively in 7.5% of households we report the use of work potential to less than 20%. Although Slovakia was below the average EU-28 very low work intensity rate within EU-28 in 2015 (while according to the EU-SILC 2015, 7.1% of the population aged 0–59 lived in very low intensity households in Slovakia (Vlačuha and Kováčová, 2016) and in the EU-28 according to Eurostat[6] there were 10.6% of such households), it should be noted that the large majority of households with very low work intensity in Slovakia did not demonstrate any work activity. In our sample, up to 92% of households with very low work intensity had zero use of labour potential throughout the whole reference period. This situation was caused by a high rate of unemployment of 11.5% in the SR in 2015 which was the 7th highest unemployment rate in the EU-28 (after Greece, Spain, Croatia, Cyprus, Portugal and Italy) while the EU-28 unemployment rate was 9.4%. We would like to note that in 2016 the unemployment rate in Europe declined substantially (by 0.9 pp in EU-28), and one of the largest drops in unemployment rate, by 1.9 pp., was recorded in Slovakia (see: Eurostat, 2017).

On the basis of a number of scientific publications (in particular those listed in the introduction to the article) and on the basis of our own experience, we have assumed that the level of household work intensity is affected by these variables observed in the EU-SILC survey: status of economic activity, the highest level of education achieved, the marital status and age of the head of household as well as the type of household, region and degree of urbanization, respectively. The population density on the territory where the household resides. For better clarity, we used custom names for variables and their variations (categories) in the analyses. Because the numbers of households in some categories were low, we have

[6] See: <http://ec.europa.eu/eurostat/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=t2020_51&language=en>.

merged them with similar categories of the relevant factor. The description of the input variables[7] and the above mentioned changes in the names and in the definition of the categories of these variables are captured in Table 2.

Table 2 Description of input explanatory variables

Original variables (EU-SILC) – categories and description		Names of new dummy variables
RB210 – Status of basic economic activity		**EAS**
1	employed	at work
2	unemployed	unemployed
3	old-age pensioner, early retirement pensioner	retired
4	other inactive person	inactive person
PE040 – The highest level of education achieved (ISCED)		**EDUCATION**
0	less than primary	Less_than_Secondary
1	primary	
2	lower secondary	
3	upper secondary	Upper_Secondary
4	post secondary (not tertiary)	Post_Secondary
5	short cycle of tertiary education	Tertiary_1
6	bachelor education	
7	master's degree or equivalent	Tertiary_2_3
8	doctoral education or its equivalent	
HT – Type of household		**HT**
5	single-person household	1adult
6	Household 2 adults, both aged 65 years	2a_0ch
7	Household of 2 adults, at least 1 at age 65+	2a_1r
8	Other households without dependent children	other_0Ch
9	Household 1 parent with at least 1 dependent child	1a_at_least_1ch
10	Household of 2 adults with 1 dependent child	2a_1ch
11	Household of 2 adults with 2 dependent children	2a_2ch
12	Household of 2 adults with 3+ dependent children	2a_at_least_3ch
13	Other households with dependent children	other_with_ch
PB190 – Marital Status		**MARITAL STATUS**
1	Single	single
2	Married	married
4	Widowed	widowed
5	Divorced	divorced
	DB100 – Degree of urbanisation	
1	region with dense population	dense
2	region with overall dense population	intermediate
3	region with sparse population	sparse

Source: EU-SILC 2015, own processing

[7] For a correct interpretation it is necessary to take into account the description of relevant variables that is stated on the website: <http://ec.europa.eu/eurostat/web/income-and-living-conditions/methodology/list-variables>.

Table 2 Description of input explanatory variables (continuation)

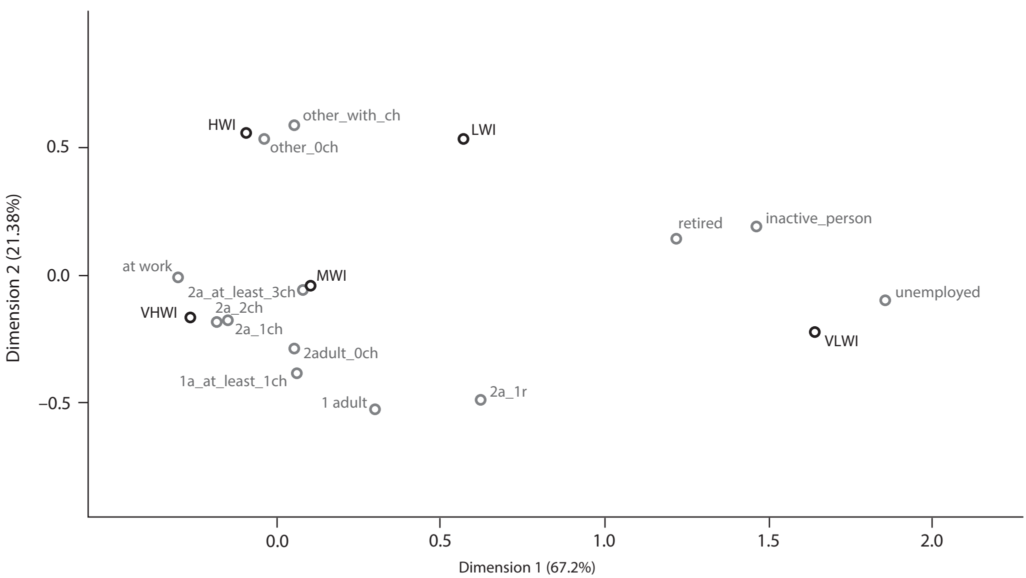
Original variables (EU-SILC) – categories and description		Names of new dummy variables
	Region	**REGION**
1	Bratislavský	BA
2	Trnavský	TT
3	Trenčiansky	TN
4	Nitriansky	NR
5	Žilinský	ZI
6	Banskobystrický	BB
7	Prešovský	PE
8	Košický	KE

Source: EU-SILC 2015, own processing

3 ASSESSMENT OF WORK INTENSITY OF SLOVAK HOUSEHOLDS WITH THE USE OF CORRESPONDENCE ANALYSIS

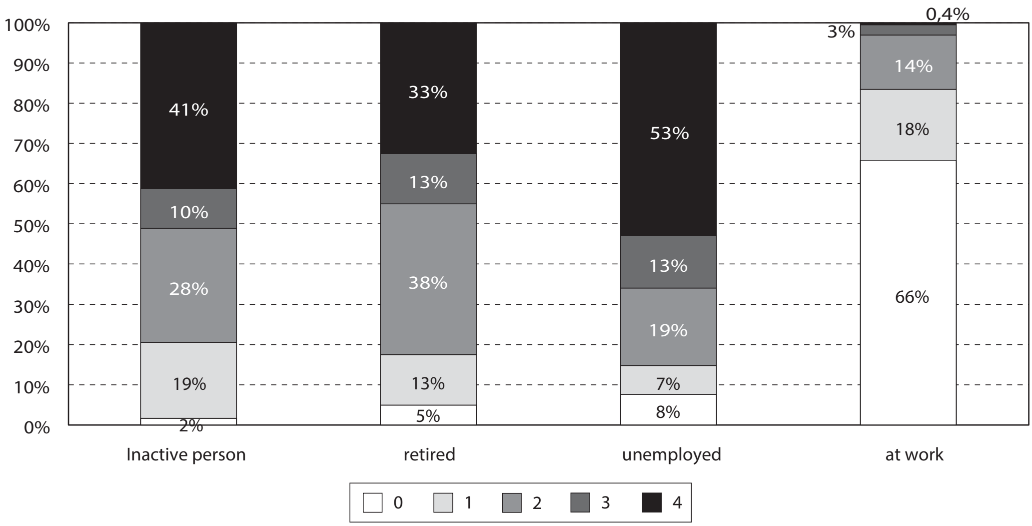
On the basis of correspondence analysis (Figure 1) and the occurrence of individual levels of work intensity of Slovak households in the sample (Figure 2), we can assert that very high labour intensity is associated especially with the households headed by the employed person. On the other hand, very low work intensity is typical for households with unemployed head of household. Correspondence analysis results confirm that "retired" and "inactive person" households in the reference year 2014 were slightly better off than households with the unemployed head. For households headed by the otherwise inactive person and for the households headed by a retired person, the very low work intensity is not as typical, and in particular, the "retired" household group is more strongly associated with low and medium work intensity compared to households with unemployed head of household.

Figure 1 Correspondence analysis of work intensity of Slovak households for factors, such as the economic activity of the household head and the household type



Source: EU-SILC 2015, customized in SAS BASE

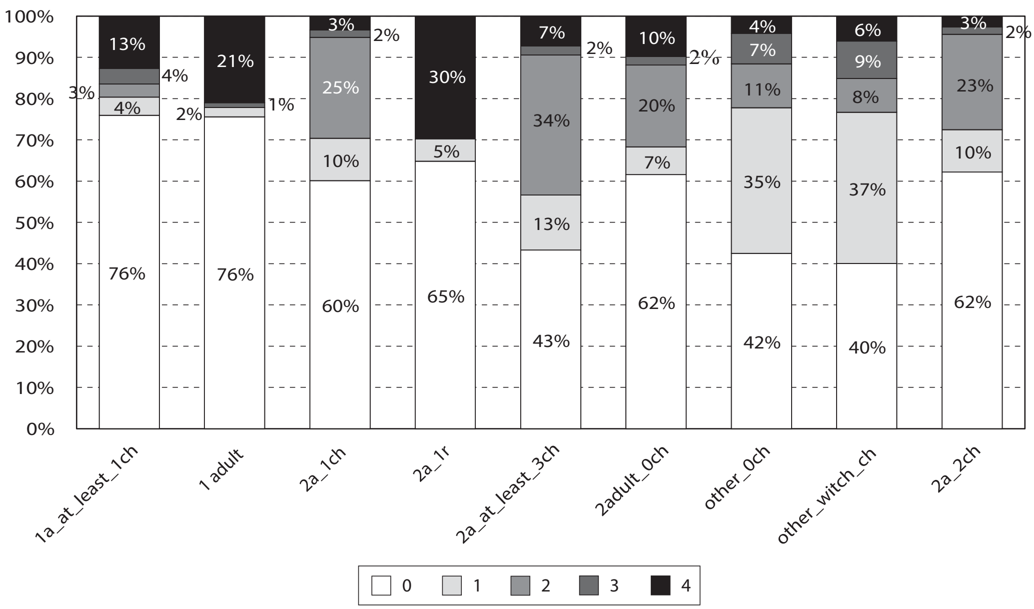
Figure 2 Household work intensity depending on the economic activity of the household head



Source: EU-SILC 2015, own processing

Correspondence analysis showed that the occurrence of very low work intensity is determined significantly more by the economic activity of the person at the head of the household as by the type of household. Very low work intensity is most typical for households where there are no more than 1 person of working age, especially for households of 2 adults, of which at least one is aged 65+ as well as single member households. A generally high representation of very low work intensity

Figure 3 Work intensity of Slovak households depending on the type of household

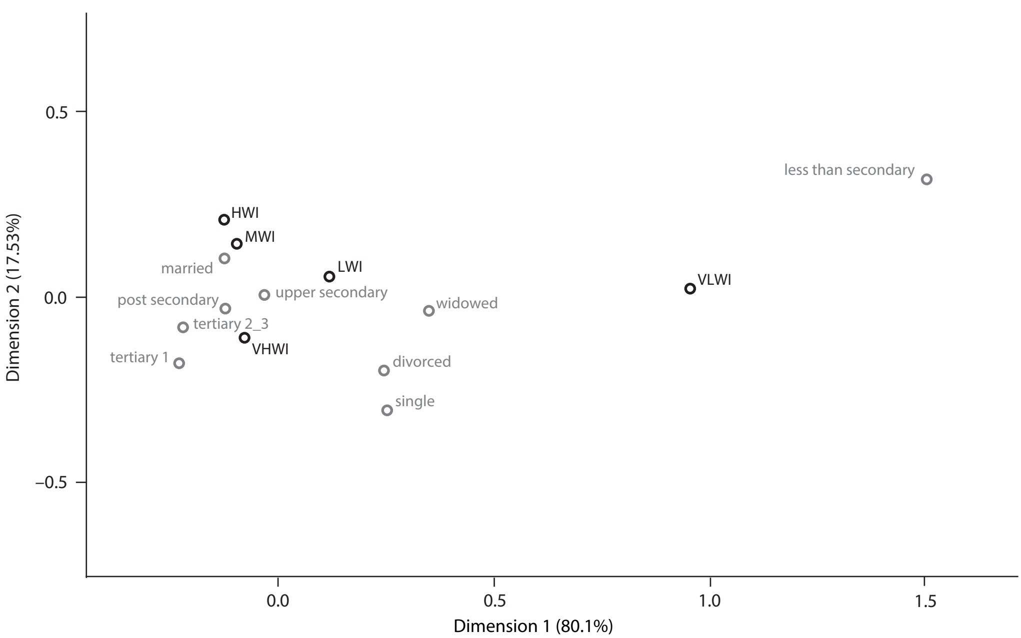


Source: EU-SILC 2015, own processing

(Figure 3) was seen, except in households of 2 adults, of which at least 1 was aged 65+ (30%) and in single member households (21%), in 1 adult household with at least 1 dependent child (13%). For "other" households, whether with or without dependent children, a high work intensity (HWI) is much more typical compared with other types of households (but not very high work intensity). Households of 2 adults and at least 3 dependent children are associated with a medium work intensity (MWI). This finding in the correspondence analysis is also confirmed by Figure 3. Generally, 2-adult households with two dependent children or one dependent child use their work potential to its best. These types of households are associated with very high work intensity (VHWI) the most and, together with households of the "other" type, are the least associated with very low work intensity (Figure 1). Moreover, in 2014, only they had the very lowest incidence of very low work intensity (under 5%) and at the same time they had a relatively high incidence of very high work intensity (Figure 3 at 60%).

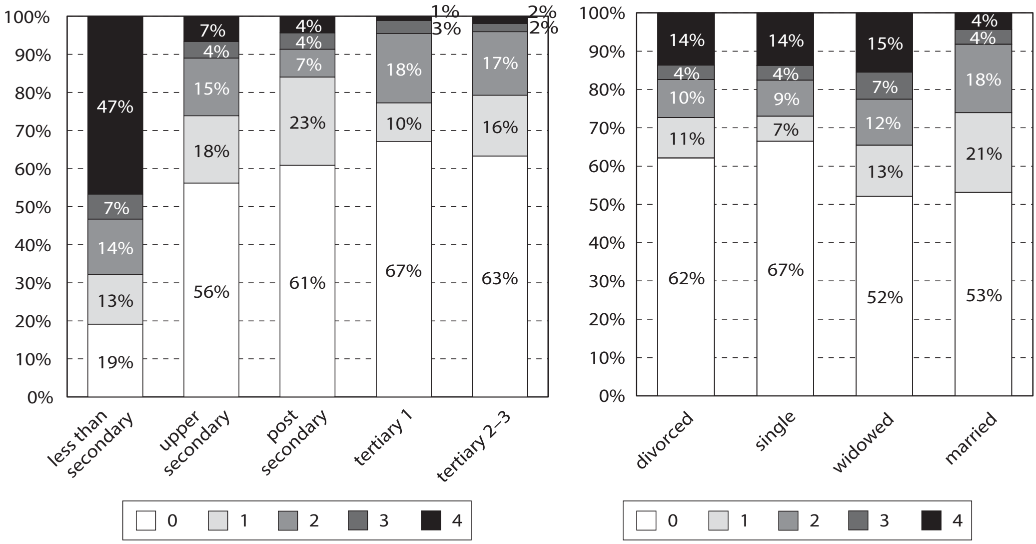
In regard to education of the person at the head of household, the results of correspondence analysis in relation to the household work intensity are very clear (Figure 4). Households headed by a person with lower than secondary education are the least of all households associated with very high work intensity, and a very low use of labour potential is typical for them. While in this group of households we recorded (Figure 5) less than ⅓ of households with high and very high work intensity, households with a person at the head with higher level of education were represented by a high and very high utilization of work potential at approximately ¾. This group of households significantly differs from other household groups, and there are no significant differences between the other groups in the use of labour potential. However, we can observe that the very low work intensity is not typical for households headed by a university graduate with either a doctorate, a master's or a doctoral degree.

Figure 4 Correspondence analysis of work intensity of Slovak households for factors, such as the highest level of education achieved and the marital status of the person at the head of the household



Source: EU-SILC 2015, customized in SAS BASE

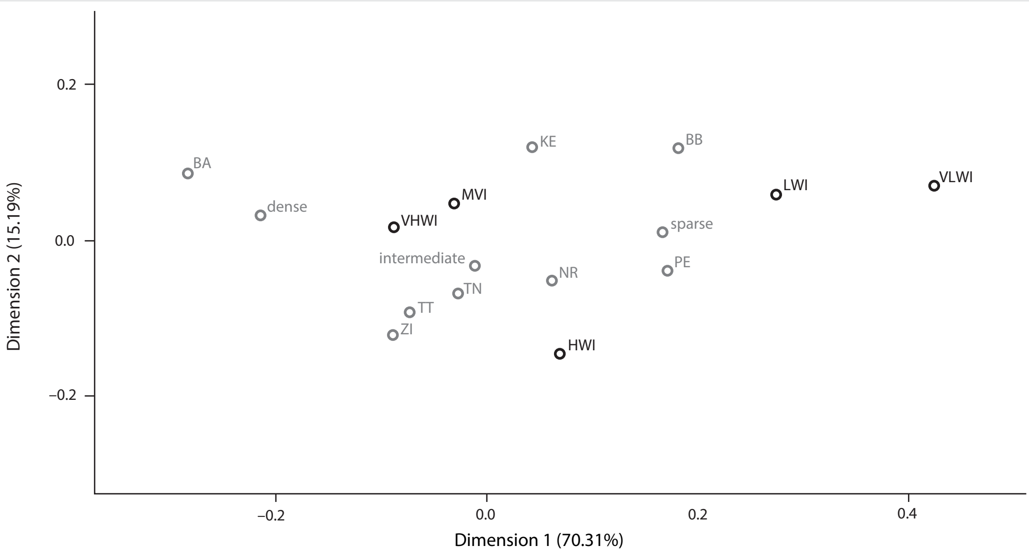
Figure 5 Work intensity of Slovak households in dependence on the highest level of education achieved and the marital status of the person at the head of the household



Source: EU-SILC 2015, own processing

The marital status of a person at the head of the household does not determine such disparities in the use of work potential as education. Very low work intensity is the least typical for households headed by a married couple. These households are more strongly associated with high and medium work intensity than other household groups. Very low and low work intensity are most typical for households with a widowed head of household.

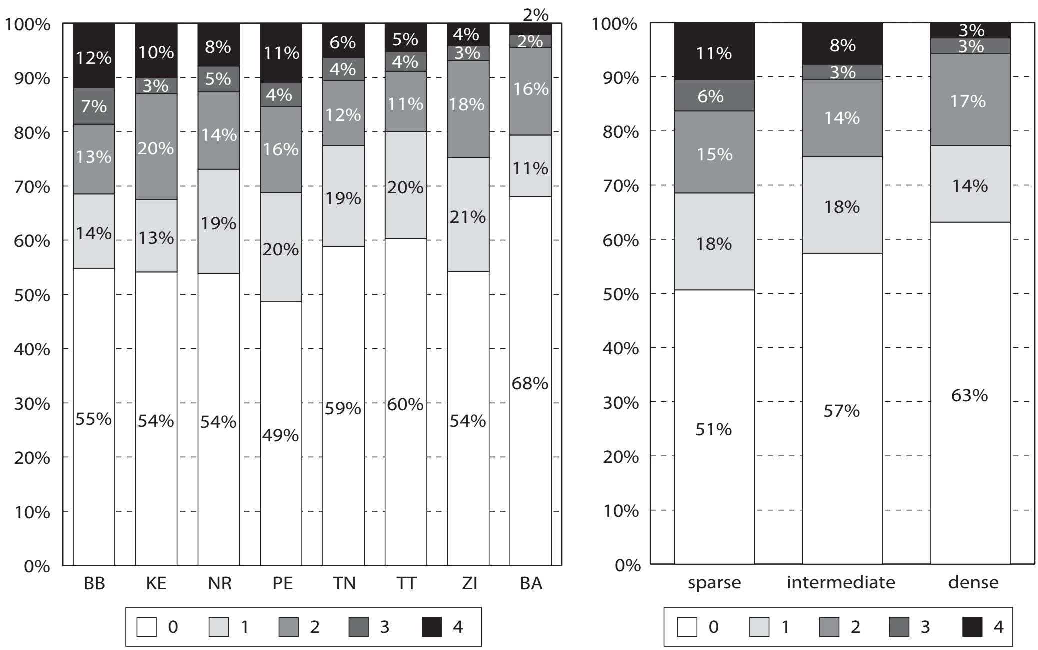
Figure 6 Correspondence analysis of work intensity of Slovak households for factors, such as region and population density on the territory where the household



Source: EU-SILC 2015, customized in SAS BASE

From a geographical point of view and in terms of population density, very low and low use of employment potential was associated with the Banská Bystrica and Prešov regions and with sparsely populated areas (Figures 6 and 7) in 2014. According to Beňuš et al. (2016), these two regions recorded one of the smallest advances in reducing the number of inhabitants living in very low work intensity households from 2010 to 2016, while the most significant decline in the very low work intensity rate was recorded in the Nitra region. Based on our analysis, the smallest threat of very low work intensity in 2014 was clearly in the Bratislava region and in households living in densely populated areas. Based on the correspondence analysis (Figure 6) and estimates of the representation of individual degrees of labour potential reduced use (Figure 7), regional disparities and discrepancies in terms of degree of urbanization are significantly higher in case of very low work intensity than in the case of very high work intensity. Of course, if we talk about relative differences.

Figure 7 Work intensity of Slovak households depending on the region and population density on the territory where the household lives



Source: EU-SILC 2015, own processing

4 ASSESSEMENT OF THE WORK INTENSITY OF SLOVAK HOUSEHOLDS USING MULTINOMIAL LOGISTIC REGRESSION

In this part of the article we will use multinomial logistic regression for the assessment of the statistical significance of the influence of the considered explanatory variables (shown in Table 2) on the degree of Slovak households work intensity. The impact of each relevant factor will be quantified by odds ratio, while the effect of other significant factors being fixed. Based on Table 3, we find that the strongest impact on work intensity is expected to be the economic activity of the person at the head of the household. This is followed by factors such as the type of household, education of the person at the head of the household and the region where the household lives. However, we quantified the smaller impact, which is still significant at the significance level of 0.05, in the case of the variable marital status of the person at the head of the household and the density of population on the territory in which the household lives.

We would like to remind that the influence of all these qualitative variables on the work intensity of Slovak households was assessed in the previous part of the article by the means of correspondence analysis. In the model of multinomial logistic regression, we also considered one continuous numeric variable – the variable age of the person at the head of the household, whose influence on work intensity is also proved to be relevant (*p-value* is less than 0.0001).

Table 3 Assessment of the multinomial logistic regression model quality and verification of the significance of the influence of considered factors on the work intensity degree of Slovak households

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	7 863.1	5 700.7
SC	7 887.4	6 428.5
-2 Log L	7 855.1	5 460.7

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2 394.5	116	<.0001
Score	2 625.9	116	<.0001
Wald	1 032.5	116	<.0001

R-Square	0.5288	**Max-rescaled R-Square**	0.5778

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
AGE	4	38.39	<.0001
EAS	12	423.68	<.0001
HT	32	421.68	<.0001
EDUCATION	20	68.78	<.0001
REGION	28	65.12	<.0001
MARITAL_STATUS	12	24.13	0.0195
URBANISATION	8	15.61	0.0483

Source: EU-SILC 2015, customized in SAS EG

In interpreting the results of the estimated logistical model, we must be aware of several facts:

- The multinomial model could be divided into several separate models, depending on the number of categories of the explanatory variable that are being examined. In our case, the dependent variable has 5 categories and the multinomial model has been divided into 4 separate models for clarity and better comparability: 1 vs. 0, 2 vs. 0, 3 vs. 0 and 4 vs. 0 (see Table 4). The values of the explanatory variables 0 to 4 represent the degree of severity of the work potential reduced use, with the value 1 representing the lowest degree of severity, that is high work intensity (HWI) and value 4 representing the highest degree of severity, that is very low work intensity (VLWI). In the above partial models 1 vs. 0 to 4 vs. 0 we compare the corresponding degree of the work potential reduced use to grade 0, which represents a very high work intensity (VHWI).
- In columns in Table 4 we present the values of the odds ratios and *p*-values for testing the significance of the relevant parameter for the logistic model. In Table 4, those *p*-values are highlighted, by which we can assume, at the significance level of 0.05, that the odds of the corresponding degree of the labour potential reduced use compared to zero degree, will be different, in the corresponding household

Table 4 Estimates of the odds ratio of a multinomial logistic regression model

Factor			1 vs. 0		2 vs. 0		3 vs. 0		4 vs. 0	
			Odds Ratio	p-value	Odds Ratio	p-value	Odds Ratio	p-value	Odds Ratio	p-value
AGE			0.983	0.0430	0.952	<.0001	0.974	0.0769	0.975	0.1206
EAS	**inactive_person**	**vs at work**	48.289	<.0001	**137.932**	<.0001	**175.831**	<.0001	**4 553.27**	<.0001
EAS	**retired**	**vs at work**	10.591	0.0077	**89.981**	<.0001	**113.047**	<.0001	**1 378.57**	<.0001
EAS	**unemployed**	**vs at work**	3.733	0.0006	**17.299**	<.0001	**42.322**	<.0001	**1 000.95**	<.0001
HT	**1a_at_least_1ch**	**vs 2a_2ch**	0.305	0.0076	**0.079**	<.0001	0.794	0.7018	2.781	0.1362
HT	**1adult**	**vs 2a_2ch**	0.168	<.0001	<0.001	0.7433	**0.159**	0.0087	**4.422**	0.0196
HT	**2a_1ch**	**vs 2a_2ch**	1.023	0.9151	1.026	0.8744	0.751	0.5584	1.343	0.5716
HT	**2a_1r**	**vs 2a_2ch**	0.591	0.4990	<0.001	0.9372	<0.001	0.9829	**11.065**	0.0287
HT	**2a_at_least_3ch**	**vs 2a_2ch**	1.893	0.0196	**2.061**	0.0006	1.799	0.3388	2.409	0.1477
HT	**2adult_0ch**	**vs 2a_2ch**	0.631	0.0849	0.814	0.3071	0.522	0.2153	1.660	0.3217
HT	**other_0ch**	**vs 2a_2ch**	5.291	<.0001	0.741	0.1897	**3.738**	0.0017	0.728	0.5471
HT	**other_with_ch**	**vs 2a_2ch**	6.038	<.0001	**0.640**	0.0337	**6.713**	<.0001	1.524	0.3801
Education	**less than secondary**	**vs tertiary 2_3**	1.530	0.2355	**2.323**	0.0212	**3.080**	0.0368	**39.726**	<.0001
Education	**post secondary**	**vs tertiary 2_3**	1.409	0.3432	0.453	0.1313	1.749	0.4321	2.770	0.2823
Education	**tertiary 1**	**vs tertiary 2_3**	0.540	0.1250	0.809	0.5335	1.105	0.8883	0.580	0.7257
Education	**upper secondary**	**vs tertiary 2_3**	0.802	0.1310	0.907	0.5043	1.187	0.6008	2.455	0.0545
REGION	**BB**	**vs BA**	1.374	0.2092	1.151	0.5713	**3.046**	0.0154	**3.824**	0.0380
REGION	**KE**	**vs BA**	1.162	0.5175	1.400	0.1113	1.095	0.8486	3.214	0.0592
REGION	**NR**	**vs BA**	1.858	0.0108	1.044	0.8640	2.073	0.1256	1.673	0.4346
REGION	**PE**	**vs BA**	1.739	0.0204	1.312	0.2563	1.689	0.2772	**4.483**	0.0201
REGION	**TN**	**vs BA**	1.633	0.0406	0.877	0.5952	2.178	0.1016	1.688	0.4544
REGION	**TT**	**vs BA**	1.896	0.0115	0.750	0.2894	1.798	0.2581	0.641	0.5491
REGION	**ZI**	**vs BA**	**2.091**	0.0017	1.524	0.0660	1.268	0.6422	1.254	0.7458
Marital_Status	**divorced**	**vs married**	1.053	0.7953	1.288	0.2258	**2.008**	0.0355	**2.480**	0.0134
Marital_Status	**single**	**vs married**	0.857	0.5448	0.739	0.2004	1.944	0.0837	0.619	0.2608
Marital_Status	**widowed**	**vs married**	1.243	0.4881	1.326	0.4153	2.295	0.0612	0.729	0.5319
URBANISATION	**intermediate**	**vs dense**	0.824	0.2288	0.866	0.3837	**0.520**	0.0472	1.637	0.2369
URBANISATION	**sparse**	**vs dense**	0.821	0.2299	1.056	0.7423	1.032	0.9188	**2.223**	0.0493

Source: EU-SILC 2015, customized in SAS EG group, from the odds in the reference household group. The household reference group is listed for the appropriate categorical variable in the name of the row – behind the word "versus" (abbreviated as "vs").

- We consider a partial model that compares households with the degree of the work potential reduced use "s" versus very high work intensity households (VHWI – level 0). If, on the basis of such a model, we find that the odds ratio for the category (households) *A* compared to level 0 is higher than for category *B* of that variable (again with respect to level 0), this does not mean that the probability of a reduced use of the work potential at the level "s" is in category *A* higher than in category *B*. We have to realize that the basis, that is the occurrence of very high work intensity, to which the comparison is made in calculation of the odds ratio, can be significantly different in categories *A* and *B*. We warn of this fact in order to avoid misinterpretation of the results.
- All estimates of regression coefficients and odds ratios calculated from them, are interpreted, assuming ceteris paribus, that is assuming that the other explanatory variables remain fixed.

According to the odds ratio for the economic activity factor, the households headed by an unemployed or otherwise inactive or retired person have at the level of significance of 0.05 significantly higher odds of high, medium, low and very low work intensity in proportion to the very high work intensity compared to households where the head of the household is employed. The probability that a household headed by an unemployed person will have a high, medium, low, or very low degree of the labour potential use compared to the probability of having a very high work intensity is approximately 4 times, 17 times, 42 times or even up to 1001 times higher (see Table 4) than in households headed by an employed person. In the case of households headed by a retired or otherwise inactive person, these odds ratios are even considerably higher in regard to households with an employed head. However, this is not caused by the fact that in these household groups there is a higher incidence of very low and low levels of labour potential use than in the households with the unemployed head, but by the fact that in these households the number of households with very high work intensity (2% and 5%) is several times lower than in households with the unemployed head (8%) (see Figure 1).

The multinomial logistic model has confirmed our previous findings that the largest threat of reduced labour potential was faced, in 2014, by two adult households, of which at least one person was 65+, as well as one-person households. Households of 2 adults, of which at least one was aged 65+, had the odds ratio of very low work intensity compared to very high work intensity, up to 11 times higher, than households of 2 adults with 2 dependent children (reference group) and even more than 15 times higher than households of the "other" type without dependent children (according to Table 4, this is the type of household with the lowest odds ratio of 4 vs. 0). Compared to the reference category (households of 2 adults with 2 dependent children), we also found statistically significantly higher odds of 4 vs. 0 in single-person households where we estimated a 4.4-fold higher chance of this unfavourable phenomenon. According to estimated shares of the very low work intensity in Figure 3, the odds ratio of 1 adult households with at least 1 dependent child have the highest risk of very low work intensity between the two mentioned household types. In this type of households, we account for approximately 13% of households with very low work intensity, which is significantly less than in the households of 2 adults, of which at least one is aged 65+ (30%) and one-person households (21%) (see Figure 3). The odds of very low utilization of the labour potential relative to the odds of very high labour potential utilization (odds ratio 4 vs. 0) was estimated to be in households of 1 adult with at least 1 dependent child 2.8 times than in households of 2 adults with 2 dependent children. However, this difference in regard to the calculated p-value (0.1362). is not statistically significant at the level of significance of 0.05. It should be emphasized that this result was significantly affected by the fact that both the occurrence of very low work intensity and the occurrence of very high work intensity were in the households of 1 adult with at least one dependent child higher than in the households of 2 adults with 2 dependent children.

As the group of households headed by a person with lower than secondary education has a significantly higher incidence of very low work intensity compared to other groups of households (Figure 5), it is not surprising that the analysis of the logistics model (Table 3 and Table 4) confirmed the significant impact of education on the risk of very low work intensity. Statistically significant odds ratios are mainly observed for a group of households headed by a person with lower than secondary education. These households have a nearly 40 times higher risk of very low work intensity relative to the odds of very high work intensity than households headed by a 2nd or 3rd educational degree graduate.

From the geographical point of view, most of the households with very high work intensity are located in the Bratislava region, which we also chose as a reference region. While in the Bratislava region, we, according to Figure 7, estimated the share of very labour intensive households at 68%, in the Prešov region it is only at 49%. The Prešov region together with Banskobystrický and Košický belongs among the regions with the highest risk of very low work intensity (see Figures 6 and 7). On the basis of the above, it is not surprising that the ratio of the probability of very low work intensity to the probability of very high work intensity is the highest in the three regions. In the Prešovský, Banskobystrický and Košický region, we determined by logistic regression that the given odds ratio is 4.5, 3.8 and 3.2, respectively times higher than in the Bratislava region. Moreover, these differences are statistically significant at the level of significance of 0.1 and, in the case of Banskobystrický and Prešovský regions, also at the significance level of 0.05. According to the odds ratio 4 vs. 0, it seems that the worst situation in terms of work potential utilization is in the households of Prešov region, which is also affected by the lowest occurrence of very high work intensity. According to Figure 7, however, the highest frequency of households with very low work intensity is in Banskobystrický region. In addition, the highest share of low work intensity households (7%) is registered in this region. The ratio of the probability of low work intensity to the probability of very high work intensity is in the region of Banská Bystrica, as in one region significantly higher than the given odds ratio (3 vs. 0) in Bratislava region, and this is approximately 3 times.

Finally, we will look at how the degree of work potential reduced use is influenced by the marital status of the head of the household and the degree of urbanization of the territory in which the household lives. While odds ratios of 1 vs. 0 – 4 vs. 0 for households with single or widowed head are not, at the significance level of 0.05, significantly different compared to the odds ratios for households headed by married person, so the households headed by the divorced person have the risk of low and the risk of very low work intensity significantly higher. Divorced households have the odds of very low work intensity in proportion to the chance of a very high work intensity almost 2.5 times higher than households headed by a married couple. Compared to single-member households and households with a widowed head, the odds ratio (4 vs. 0) for households with a divorced head is 4.0 times and 3.4 times higher, respectively.

Although the degree of urbanization does not have such a large impact on the threat of labour potential reduced use, in the case of this factor a category arose for which an increased risk of very low work intensity is typical. These are households living in a sparsely populated area that have a 2.2-fold higher probability of occurrence of very low work intensity compared to the likelihood of very high work intensity as households living in densely populated areas where the threat of very low work intensity is the lowest.

CONCLUSION

In Slovakia in 2014, their work potential was best used by 2 adult households with 2 dependent children or 1 dependent child, confirming the results of the correspondence analysis as well as the logistic regression presented in this article. This is due to the fact that these types of households experienced the lowest incidence of very low work intensity and a relatively high proportion of households with very high work intensity. It is not surprising that, in terms of economic activity and the highest level of education, a very high work intensity was typical for those households headed by an employed person, a person with a master's or a doctoral degree of education. Although other factors did not cause such large disparities in the use of labour potential as the above-mentioned factors, they also significantly determined the level of work intensity of Slovak households. The best use of their working potential was in the households headed by a married couple, households from the Bratislava region and households living in densely populated areas. In these three groups of households we observed a very low work intensity below 5%.

Very low work intensity in 2014 was most typical for households with a maximum of 1 person of working age, especially for households of 2 adults, of which at least 1 is aged 65+ and single households. Due to the high risk of very low work intensity, we assess these two types of households as households with the worst use of labour potential, despite the fact that, besides the high incidence of very low work intensity, there was a high incidence of very high work intensity (with negligible representation of other levels of work intensity). All applied statistical methods have clearly demonstrated that the type of household is a relevant factor determining the intensity of work of Slovak households. But we have to say that economic activity and education have had a greater impact on the use of labour potential. For households headed by an unemployed person, we have quantified that the odds ratio of the very low work intensity in regard to the very high work intensity is up to 1000 times higher than in the households with an employed head. Our analyses have confirmed that with increasing level of education, the use of households' work potential is improving and that the greatest threat of very low work intensity is in households headed by a person with lower than secondary education. The odds ratio of very low work intensity in regard to very high work intensity we estimated for these households is almost 40 times the odds ratio quantified for households that are the least threatened by exclusion from the labour market (households headed by a university graduate with the second or third degree of higher education). Other factors did not determine such large differences in the odds ratio of the very low work intensity and the very high work intensity between individual household groups. Correspondence analysis and analysis of the multinomial logistics model, however, confirmed that even in terms of the marital status of the person at the head of the household and in terms of the region and the density of settlement of the territory in which the household lives, there were significant differences in the use of labour potential in Slovakia in 2014. In regard to these three factors, the households which were using their work potential at its lowest, were the households headed by a divorced person, households from Prešov and Banská Bystrica regions and households living in sparsely populated areas.

ACKNOWLEDGMENT

The paper was supported by a grant agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic VEGA. Project: VEGA no. 1/0548/16 *The progress of the SR in achieving of the strategy Europe 2020 in the area of poverty and social exclusion.*

References

ALLISON, P. D. *Logistic Regression using SAS. Theory and Application*. 2nd Ed. North Carolina, USA: SAS Institute, 2012.

AYLLÓN, S. AND GÁBOS, A. *The interrelationships between the Europe 2020 social inclusion indicators* (No. 15/01) [online]. Herman Deleeck Centre for Social Policy, University of Antwerp, 2015. [cit. 7.6.2017] <http://www.centrumvoorsociaalbeleid.be/ImPRovE/Working%20Papers/ImPRovE%20WP%201501_1.pdf>.

BEŇUŠ, O., KOVÁČIK, M., ŽUFFOVÁ, E. Measuring Development of Selected Poverty Risk Indicators in V4 Countries with Specific Focus on Slovak Republic and its Regions. *Acta Regionalia et Environmentalica*, 2016, 13(1), pp. 22–26.

EUROSTAT. <http://ec.europa.eu/eurostat/web/income-and-living-conditions/methodology/list-variables>.

EUROSTAT. <http://ec.europa.eu/eurostat/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=t2020_51&language=en>.

EUROSTAT. <http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_definition_of_dimensions>.

EUROSTAT. *Unemployment statistics* [online]. 2017. [cit. 3.8.2017] <http://ec.europa.eu/eurostat/statistics-explained/index.php/Unemployment_statistics>.

GUAGNANO, G., SANTARELLI, E., SANTINI, I. *Subjective poverty in Europe: the role of household socioeconomic characteristics and social capital* (No. 113/13) [online]. SAPIENZA University of Rome, 2013. [cit. 10.6.2017] <https://www.memotef.uniroma1.it/sites/dipartimento/files/wpapers/documenti/FullTextWP113.pdf>.

HICK, R. AND LANAU, A. *In-work poverty in the UK: Problem, policy analysis and platform for action* [online]. Final Report, Cardiff University, 2017. [cit. 7.6.2017] <https://orca.cf.ac.uk/103013/1/Hick%20and%20Lanau%20_%20In-Work%20Poverty%20in%20the%20UK.pdf>.

KIS, B. A. AND GÁBOS, A. Consistent poverty across the EU. *Corvinus Journal of Sociology and Social Policy*, 2016, 7(2), pp. 3–27.

LOGUE, C. AND CALLAN, T. *Low Pay, Minimum Wages and Household Incomes: Evidence for Ireland* [online]. Dublin: Economic and Social Research Institute, 2016. [cit. 3.7.2017] <https://www.esri.ie/pubs/BP201703.pdf>.

MELOUN, M. AND MILITKÝ, J. *Kompendium statistického zpracování dat*. 3rd Revised Ed. Prague: Karolinum, 2012.

MELOUN, M. AND MILITKÝ, J. *Interaktivní statistická analýza dat*. 3rd Revised Ed. Prague: Karolinum, 2012.

MONTE, A., SCHOIER, G., ZUCCHET, I. An Evaluation of the Dynamics of Economic Well-Being in Europe: A Statistical Analysis. *Rivista Italiana di Economia Demografia e Statistica*, 2017, 71(1), pp. 101–112.

MYSÍKOVA, M., VEČERNÍK, J., ŽELINSKÝ, T. Impact of the Low Work Intensity on Poverty in the Czech Republic and the Slovak Republic. *Journal of Economics*, 2015, 63(6), pp. 555–575.

RAMON, C. L., STROUP, W. W., FREUND, R. J. *SAS for Linear Models*. 4th Revised Ed. North Carolina, USA: SAS Institute, 2010.

RASTRIGINA, O., LEVENTI, C., SUTHERLAND, H. *Nowcasting risk of poverty and low work intensity in Europe* (No. EM9/15) [online]. EUROMOD Working Paper. Colchester: ISER, University of Essex: 2015. [cit. 12.6.2017] <https://www.econstor.eu/bitstream/10419/113340/1/828386390.pdf>.

REDMOND, P. *Very Low Work Intensity in Ireland and the EU* [online]. 2016. [cit. 24.6.2017] <http://www.publicpolicy.ie/wp-content/uploads/Redmond2016.pdf>.

ŘEZANKOVÁ, H. *Analýza dat* z *dotazníkových* šetření. Praha: Professional Publishing, 2007.

STANKOVIČOVÁ, I. AND VOJTKOVÁ. M. *Viacrozmerné* štatistické *metódy* s *aplikáciami.* Bratislava: Iura Edition, 2007.

STATSOFT. *STATISTICA Formula Guide: Logistic Regression* [online]. 2013. [cit. 9.8.2017] <http://documentation.statsoft.com/portals/0/formula%20guide/Logistic%20Regression%20Formula%20Guide.pdf>.

ŠOLTÉS, E. AND ŠOLTÉSOVÁ, T. Occurrence of low work intensity in Slovakia in relation to assessment of poverty and social exclusion. In: *The 10th Professor Aleksander Zelias: proceedings [of] international conference on modelling and forecasting of socio-economic phenomena,* Zakopane: Polska Akademia Nauk, 2016, pp. 189–198.

VLAČUHA, R. AND KOVÁČOVÁ, Y. *EU SILC 2015. Indikátory chudoby* a *sociálneho vylúčenia* [online]. Bratislava: Štatistický úrad Slovenskej republiky, 2016. [cit. 20.6.2017] <http://slovak.statistics.sk>.

WARD, T. AND OZDEMIR, E. *Measuring low work intensity–an analysis of the indicator* [online]. ImPRovE Poverty, Discussion Paper No. 13/09, Antwerp, 2013. [cit. 20.6.2017] <http://www.centrumvoorsociaalbeleid.be/ImPRovE/Working%20Papers/ImPRovE%20WP%201309_1.pdf>.

WATSON, D., MAÎTRE, B., WHELAN, C. T. *Work and Poverty in Ireland: An Analysis of CSO Survey on Income and Living Conditions 2004–2010* [online]. Dublin: Economic and Social Research Institute, 2012. [cit. 3.7.2017] <http://www.esri.ie/pubs/BKMNEXT226.pdf>.

WHELAN, C. T. AND MAÎTRE, B. The Great Recession and the changing distribution of economic vulnerability by social class: The Irish case. *Journal of European Social Policy*, 2014, 24(5), pp. 470–485.