Analysis of work intensity in Slovakia using testing and estimation of linear combinations of GLM parameters

Šoltés, Erik; Komara, Silvia; Šoltésová, Tatiana; Mišút, Martin

ARGUMENTA OECONOMICA
No 1 (50) 2023
ISSN 1233-5835; e-ISSN 2720-5088

Erik Šoltés, Silvia Komara**, Tatiana Šoltésová***, Martin Mišút*****

Analysis of work intensity in Slovakia using testing and estimation of linear combinations of GLM parameters

Not only unemployment itself but also the reduced work intensity of a household has a major impact on the social exclusion of a person. The work intensity of households is currently being monitored in Europe mainly for purposes of identifying those people or households that are excluded from the labour market. The households' work intensity directly affects the inclusion or exclusion from the labour market, which is one of the three social exclusion dimensions. Moreover, it also, as confirmed by several studies, fundamentally affects the other two dimensions of social exclusion, namely income poverty and material deprivation. The aim of the paper was to assess which factors in interaction with the economic activity status of a person significantly affect the household's work intensity and, depending on these factors, to estimate the household's work intensity. For this purpose, the general linear model and the associated analysis of marginal means and the contrast analysis were used. The analyses are based on a database EU-SILC 2020 for the Slovak Republic and performed in the SAS Enterprise Guide and by means of PROC GLM in the SAS programming language using CONTRAST and ESTIMATE statements. The article examines between which levels of significant factors there is a significant difference in terms of a household's work intensity and in particular provides estimates of work intensity depending on the household type, educational attainment level and the age of a person. At the same time, in all three cases households are broken down by the economic activity status of the person. The presented analyses revealed categories of persons that are the most and the least threatened by labour market exclusion from the point of view of the considered factors.

Keywords: work intensity, exclusion from the labour market, poverty and social exclusion, general linear model (GLM), least square means, contrast analysis

JEL Classification: C12, C21, E24
DOI: 10.15611/aoe.2023.1.03

Quote as: Šoltés, E., Komara, S., Šoltésová, T., Mišút, M. (2023). Analysis of work intensity in Slovakia using testing and estimation of linear combinations of GLM parameters. *Argumenta Oeconomica*, 1(50), s. 43-66.

* Department of Statistics, University of Economics in Bratislava, Slovakia. ORCID: 0000-0001-8570-6536.
** Department of Statistics, University of Economics in Bratislava, Slovakia. ORCID: 0000-0001-6641-7456.
*** Department of Mathematics and Actuarial Science, University of Economics in Bratislava, Slovakia. ORCID: 0000-0002-0953-2519.
**** Department of Applied Informatics, University of Economics in Bratislava, Slovakia. ORCID: 0000-0002-5545-2624.

1. Introduction

The work intensity (WI) of households is an indicator that is monitored within the framework of sustainable development strategies in the EU, e.g. in the Europe 2030 strategy. Households with very low work intensity are referred to as (quasi-)jobless households and members of such households are considered to be excluded from the labour market. The active participation in the labour market plays an important role in the fight against poverty and social exclusion, hence (quasi)joblessness is one of the three components of the composite indicator AROPE (at risk of poverty or social exclusion), which monitors the EU's progress on social inclusion. Although in 2019 the share of persons living in (quasi-)jobless households in Slovakia (4.8%) was slightly lower than in the EU (6.1%), it is important to identify factors that increase the risk of exclusion from the labour market, also regarding the unfavourable events in the early 2020s. It is a realistic assumption that the COVID-19 pandemic, the energy crisis, inflation, and the war in Ukraine will increase the share of the population in Slovakia (as in other EU countries) that will have to face social exclusion (including exclusion from the labour market).

The above-mentioned facts and presumptions motivated the author to analyse the work intensity in the Slovak Republic. This paper is not limited only to the very low work intensity but focuses on the labour intensity index. This index is a target continuous numerical variable, which was analysed depending on various factors via the analysis of marginal means and contrast analysis, which are based on a general linear model. The aim of the paper was to assess the influence of the most fundamental socio-economic and socio-demographic factors on work intensity, while assessing the influence of other relevant factors. The following research tasks are also oriented on the most relevant factors:

- to assess whether the impact of factors on work intensity is different or the same for different statuses of economic activity,
- for each factor, identify categories between which there are no significant differences and identify those categories or clusters of categories between which there are demonstrable differences in work intensity,
- to quantify the mean of work intensity for individual groups of persons and identify risk groups of persons in terms of exclusion from the labour market.

2. Literature review

Exclusion from the labour market significantly increases the risk of material deprivation and income poverty. De Graaf-Zijl and Nolan (2011) stated that the dependence between these three dimensions does not have a consistent pattern in groups of countries classified together in terms of welfare regime or geographically. García-Gómez et al. (2021) found that this dependence increased significantly in the countries most affected by the economic crisis in the period 2008-2014. Duiella and

Turrini (2014) came to a similar conclusion, and also identified a positive relationship among them, which became stronger after 2010 in countries most severely hit by the crisis. Based on the above, it was supposed that this relationship will also intensify due to the COVID-19 pandemic and the deteriorating economic situation caused by the energy crisis, inflation, and the war in Ukraine. Verbunt and Guio (2019) confirmed that work intensity is very effective in explaining within-country differences in the risk of income poverty/material deprivation in some CEE countries (including Slovakia). In Slovakia, after the period of financial and economic crisis, unemployed persons living in households with a high and medium level of work intensity had markedly higher chances to move to employment, compared to the unemployed in households with low work intensity (Gerbery and Miklošovic, 2020). Fabrizi and Mussida (2020) found that living in a work-poor household is associated with living in consistent poverty (people at consistent poverty are those who are both at risk of poverty and simultaneously experiencing enforced deprivation).

In addition, labour market exclusion also has a negative impact on the population of children and young people and on their social exclusion in the future. Guio and Vandenbroucke (2019) stated that (quasi)joblessness is an important driver of child deprivation in Belgium, even when income is controlled for. Analyses by Cantó et al. (2022) revealed that other household members' employment levels and economic difficulties have strong effects on youth economic outcomes.

This paper does not deal only with low (or very low) work intensity, as other degrees of work intensity can also be associated with poverty and social exclusion. For example, Kis and Gábos (2016) showed that in the new member states of the EU not only low and very low household work intensity is positively associated with a higher risk of consistent poverty, but also medium work intensity. Naturally, higher work intensity is positively correlated with social inclusion. Fabrizi and Mussida (2020) showed that higher work intensity of Italian households with dependent children significantly reduces the probability of falling into poverty and social exclusion. Although having a job is not a sufficient condition to avoid poverty, either in terms of (monetary) objective or subjective poverty (Filandri et al., 2020), low work intensity is a crucial micro-determinant of in-work poverty. Hick and Lanau (2017) stated that work intensity of the household is a very strong predictor of in-work poverty. Colombarolli (2021) confirmed that work intensity is negatively associated with in-work poverty, but the relation is stronger with the objective in-work poverty rather than subjective one. Research of work intensity is therefore of great importance also in terms of in-work poverty, the prevention of which is very important for raising living standards and ensuring its convergence in the EU member states.

Poverty and social exclusion analyses use statistical modeling intensively, using different types of generalised linear models. Among the most commonly used are binomial logit models applied by e.g. Ćwiek and Ulman (2019), Šoltés and Ulman (2015), Mysíková et al. (2019), and multinomial logistic models applied by e.g.

Sanchez-Sellero and Garcia-Carro (2020), Calegari et al. (2021), Verbunt and Guio (2019). This study used a general linear model, which is also a special case of a generalised linear model. Unlike other analyses in the field of poverty and social exclusion, the paper focuses on the analysis of marginal means and the contrast analysis.

In selecting the explanatory factors, the author relied on the results of the previous research and the work of other researchers. These factors include the status of economic activity, education, type of household, age, marital status, health condition, region, and degree of urbanisation. The study focused especially on the impact of the first four factors on work intensity, whose significant impact on poverty and social exclusion was confirmed by, among others, Nieuwenhuis and Maldonado (2018) and Peña-Casas et al. (2019).

3. Method

This paper proceeded from the general linear model GLM (Littell et al., 2010), based on which the influence of categorical factors and their interactions on a continuous numerical response variable characterising work intensity were assessed. In terms of interpreting the results, it is important to note that in the research used factors with fixed effects (Searle and Gruber, 2017), and for categorical factors, the author used indicator (dummy) coding (Darlington and Hayes, 2016). The interaction was based on the crossed classification structure (Littell et al., 2010).

The general linear model can be written in matrix form as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{1}$$

Matrix $\mathbf{X}$ is not full rank, and a generalised inverse method is used to estimate the vector of parameters $\boldsymbol{\beta}$, the result of which is an estimate

$$\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-}\mathbf{X}^{\mathrm{T}}\mathbf{y}, \tag{2}$$

where matrix $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-}$ is a generalized inverse matrix that must satisfy at least the first of the Penrose conditions (Searle and Gruber, 2017). The estimation of the vector of parameters $\boldsymbol{\beta}$ obtained by the generalised inverse method is not unique, but there is a group of linear functions of the parameters, referred to as estimable functions $\mathbf{L}\boldsymbol{\beta}$ (Elswick et al., 1991), for which there is a single solution (for more detail see e.g. Agresti, 2015 and Littell et al., 2010).

As the aim of the paper was, among other things, to assess between which categories of relevant factors there is a significant difference, the subject of interest was the testing of general linear hypotheses. To test the general linear hypothesis H_0: $\mathbf{L}\boldsymbol{\beta} = \mathbf{0}$ (cf. McFarquhar, 2016, Poline et al., 2007, Searle and Gruber, 2017), the following test statistic was used

$$F = \frac{\dfrac{(\mathbf{Lb})^{\mathrm{T}} \cdot \left[\mathbf{L}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-} \mathbf{L}^{\mathrm{T}}\right]^{-1} \cdot (\mathbf{Lb})}{l}}{\dfrac{SSE}{n-p}}, \tag{3}$$

where l is the number of independent rows of matrix $\mathbf{L}$, SSE is the sum of the squared residuals, n is the sample size, and p is the number of parameters of the GLM. Thus the null hypothesis is rejected if the value of the test statistic satisfies the inequality

$$F > F_{1-\alpha}(l;\, n-p). \tag{4}$$

The above test was used to test simple hypotheses (if $l = 1$) and to simultaneously test multiple hypotheses (if $l \geq 2$). To test simple hypotheses, a t-test was also used, or alternatively, an interval estimate was also constructed (cf. Kuznetsova et al., 2017, Littell et al., 2010, and Westfall and Tobias, 2007).

The analyses presented in the paper were based on unbalanced data, while the author assessed the impact of several effects. In such a situation, group arithmetic means do not provide an adequate picture of the response of the target variable for the particular factor because they do not take into account other effects, which may lead to the Simpson paradox (Wang et al., 2018). Cai (2014) stated that if the data are unbalanced, arithmetic means are not appropriate because they do not consider that not all factors have the same chance of influencing the target variable. In such cases, it is appropriate to estimate the marginal means, which are based on the model (in this case on the GLM). The marginal mean is also referred to as the LS-mean (*Least Squares mean*; Goodnight and Harvey, 1997) or the EM-mean (*Estimated Marginal mean*; Searle et al., 1980). The estimated marginal means or least squares means are predicted means that are calculated from the fitted model and are adjusted appropriately for any other variable (Suzuki et al., 2019).

This study employed marginal mean analysis using the LSMEANS statement and contrast analysis (Dean et al., 2017, Kim and Timm, 2006, Schad et al., 2020) using the CONTRAST and ESTIMATE statements within PROC GLM (SAS Institute Inc., 2018b) and PROC GENMOD (SAS Institute Inc., 2018a) in the SAS programming language. The procedures in SAS presented in this paper are largely universal and are also used in other software and open-source systems (Lenth, 2016, Tabachnick and Fidell, 2013).

4. Database

The author analysed work intensity (WI) in Slovakia using a general linear model with explanatory variables listed in Table 1. The analyses were based on the EU-SILC 2020 database (with the reference year 2019) provided by the Statistical Office of the Slovak Republic. The statistical unit is the person to whom the WI of the household in which this person lives is assigned.

Table 1

Description of input explanatory variables

Original variables (EU-SILC) – categories and description	**Names of new dummy variables**
1	2
RB210 – Economic activity status	**EAS**
Other inactive person	IP
Unemployed person	UP
Employed person	EP
HT – Household type	**HT**
Single-person household	1A_0Ch
Single parent household with at least 1 dependent child	1A_1^{+}Ch
2-adult household, at least 1 aged 65+	2A(1^{+}R)
2-adult household without dependent children	2A_0Ch
2-adult household with 1 dependent child	2A_1Ch
2-adult household with 3+ dependent children	2A_3^{+}Ch
Other households without dependent children	Other_0Ch
Other households with dependent children	Other_1^{+}Ch
2-adult household with 2 dependent children	2A_2Ch
PE040 – The highest level of education achieved (ISCED)	**Education**
Pre-primary (ISCED 0)	ISCED 0-2
Primary (ISCED 1)	
Lower secondary (ISCED 2)	
Upper secondary (ISCED 3)	ISCED 3-5
Post-secondary (not tertiary) (ISCED 4)	
Short cycle of tertiary education (ISCED 5)	
Bachelor or equivalent (ISCED 6)	ISCED 6-8
Master‘s or equivalent (ISCED 7)	
Doctorate or equivalent (ISCED 8)	
RX010 – Age	**Age**
Age at the end of income reference period	<30
	30-40
	40-50
	50-60
PH010 – General health	**Health**
Very bad	Bad
Bad	
Fair	Fair

1	2
Good	Good
Very good	
PB190 – Marital Status	**Marital status**
Single	Single
Married	Married
Widowed	Widowed
Divorced	Divorced
Region	**Region**
Banská Bystrica	BB
Prešov	PO
Košice	KE
Žilina	ZA
Trenčín	TN
Trnava	TT
Nitra	NR
Bratislava	BA
DB100 – Degree of urbanisation	**Urbanisation**
Thinly populated area	Sparse
Intermediate area	intermediate
Densely populated area	Dense

Source: own elaboration based on EU-SILC data.

The definition of the target variable WI is given by the methodology used by Eurostat to monitor exclusion from the labour market. For the EU-SILC 2020 database, the definition used in the Europe 2020 Strategy was applied. Based on this, the household work intensity was defined as the proportion of the total number of months during which in the course of the income reference year all members of the productive-age household worked, and the total number of months that the same household members could theoretically work, under state legislation, during the same period. A person of productive age means a person aged 18-59 with the exclusion of students in the 18-24 age group. Note that from 2021 onwards, a modified definition was applied, as used in the Europe 2030 Strategy (Eurostat, 2022). As by definition, the WI is not assigned to some persons, information from 7,424 persons was entered in the analysis, although 13,800 persons were included in the EU-SILC 2020 survey.

In the analyses, a continuous numerical variable WI was applied in the sense of the above definition, but also interpreting the results of the analyses with respect to the degrees of work intensity. According to the Eurostat methodology, for work intensity of households from intervals $\langle 0\%; 20\%)$, $\langle 20\%; 45\%)$, $\langle 45\%; 55\%\rangle$, $(55\%; 85\%\rangle$, and $(85\%; 100\%\rangle$, the degrees of very low work intensity (VLWI), low work intensity (LWI), medium work intensity (MWI), high work intensity (HWI), and very high work intensity (VHWI), respectively, were assigned.

5. General linear model for work intensity

5.1. Regressor selection

Using the stepwise regression method (Agresti, 2015), the regressors listed in Table 2 were included in the model. Naturally, the WI is fundamentally influenced by economic activity. The EAS variable alone explains more than 1/4 of the WI variability (i.e. 28.09%). The impact of the other variables listed in Table 1 (excluding the urbanisation variable) also proved to be significant. Originally, these variables were considered separately (not in interaction), in which case the model explained the WI variability to about one-third. In fact, the above variables affect the WI differently for different statuses of economic activity, which confirms the significance of the individual interactions (Table 2). Thanks to the consideration of interactions, it was possible to substantially increase the explained variability of the WI to more than 50% (i.e. 50.45%).

Table 2

Verification of statistical significance of the model and the influence of factors on WI

Source	DF	Sum of squares	Mean square	F value	Pr > F
Model	77	309.633	4.021	97.12	<.0001
Error	7 346	304.163	0.041		
Corrected total	7 423	613.796			

R-square	Coeff var	Root MSE	WI mean
0.5045	25.121	0.203	0.810

Source	DF	Partial R-square	Model R-square	Type III SS	Mean square	F value	Pr > F
EAS	2	0.2809	0.2809	33.113	16.556	399.86	<.0001
EAS*HT	24	0.1271	0.4080	19.040	0.793	19.16	<.0001
EAS*AGE	9	0.0462	0.4543	8.979	0.998	24.10	<.0001
EAS*EDUCATION	6	0.0267	0.4809	13.564	2.261	54.60	<.0001
EAS*REGION	21	0.0120	0.4929	6.657	0.317	7.66	<.0001
EAS*HEALTH	6	0.0062	0.4991	4.043	0.674	16.27	<.0001
EAS*MS	9	0.0053	0.5045	3.258	0.362	8.74	<.0001

Source: EU-SILC 2020, own processing in SAS EG.

In Table 2, the regressors are ranked/classified/sorted according to their contribution to explaining the variability of the WI variable. After the EAS variable, the WI is most strongly influenced by household type (HT), age and education, which in interaction with economic activity contributed to the explanation of the WI variability at 12.71%, 4.62%, and 2.67%, respectively. The other three interactions had a significant effect on the WI, but their contribution to the explanation of the WI variability was less than 2%.

This paper, in addition to the impact of economic activity itself, focused on quantifying the impact of the other three most important factors, namely the type of household, age and education.

5.2. Analysis of LS means and contrast analysis in GLM

In this section, the author applied the analysis of marginal means, in order to find out between which pairs of categories of individual factors (HT, age, and education) there is a significant difference in terms of WI, when assessing the influence of other factors. Given the interaction of these factors with economic activity, this comparison was made separately for each economic activity. Contrast analysis served to create clusters of several categories of the relevant factor, so that from the WI perspective, there was no significant difference between the categories belonging to the cluster, and at the same time there was a demonstrable difference between the clusters. For such clusters of categories, the study estimated the mean of WI, providing a picture of the impact of individual factors on the WI and allowing to identify the most vulnerable persons in terms of exclusion from the labour market.

5.2.1. Analysis of the impact of interaction *EAS*×*HT*

Since the type of household has the greatest influence on the WI of the assessed factors, it was possible to illustrate the procedure with the example of this factor. In Table 3, there are estimated LS-means of the WI for individual types of households, especially for three statuses of economic activity (since the WI is assigned only to persons under 60 years of age, the status of economic activity Retired was not considered).

The p-values matrices for LS-means equality tests show between which types of households there is no significant difference from the WI point of view. Originally, it was one matrix, which, due to its size, was divided into three submatrices belonging to the individual statuses of economic activity. Therefore, the author did not report test results between individual statuses, but confirmed that the WI means for employees (EP) were significantly higher ($p < 0.0001$) than for the other two statuses, and this applied to all types of households. Next, the study looked at the influence of HT only within the individual statuses of economic activity. While for inactive persons (IP), the differences were insignificant only between pairs of household types, in the case of the other two statuses of economic activity, the similarity

Table 3

Comparison of LS-means of WI for effect $EAS \times HT$

EAS = IP			**Least Squares means for effect EAS*HT Pr > |t| for H_0: LSMean(i) = LSMean(j) Dependent variable: WI (EAS = IP)**									
HT	**WI LSMEAN**	**i**	**i/j**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1A_0Ch	0.0542	1	1		<.0001	**0.5901**	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
1A_1+Ch	0.5560	2	2	<.0001		<.0001	<.0001	0.0065	0.0011	<.0001	<.0001	**0.7009**
2A(1+R)	0.0266	3	3	**0.5901**	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
2A_0Ch	0.2936	4	4	<.0001	<.0001	<.0001		<.0001	<.0001	**0.0538**	<.0001	<.0001
2A_1Ch	0.4881	5	5	<.0001	0.0065	<.0001	<.0001		0.2628	<.0001	0.0022	<.0001
2A_3+Ch	0.4618	6	6	<.0001	0.0011	<.0001	<.0001	**0.2628**		<.0001	**0.2605**	<.0001
Other_0Ch	0.3445	7	7	<.0001	<.0001	<.0001	**0.0538**	<.0001	<.0001		<.0001	<.0001
Other_1+Ch	0.4373	8	8	<.0001	<.0001	<.0001	<.0001	0.0022	**0.2605**	<.0001		<.0001
z_2A_2Ch	0.5653	9	9	<.0001	**0.7009**	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

EAS = UP			**Least Squares means for effect EAS*HT Pr > |t| for H_0: LSMean(i) = LSMean(j) Dependent variable: WI (EAS = UP)**									
HT	**WI LSMEAN**	**i**	**i/j**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1A_0Ch	0.4063	1	1		0.0049	<.0001	**0.9565**	**0.6438**	0.3735	0.4153	0.7073	0.0826
1A_1+Ch	0.0475	2	2	0.0049		**0.4108**	0.0035	0.0020	0.0015	0.0011	0.0021	0.0003
2A(1+R)	0.1531	3	3	<.0001	**0.4108**		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
2A_0Ch	0.4091	4	4	**0.9565**	0.0035	<.0001		**0.5994**	**0.3437**	**0.2903**	**0.6606**	0.0363
2A_1Ch	0.4313	5	5	**0.6438**	0.0020	<.0001	**0.5994**		0.5557	0.7248	0.8369	0.0929
2A_3+Ch	0.4676	6	6	**0.3735**	0.0015	<.0001	**0.3437**	**0.5557**		0.6860	0.4389	0.5067
Other_0Ch	0.4442	7	7	**0.4153**	0.0011	<.0001	**0.2903**	**0.7248**	**0.6860**		0.4460	0.1274
Other_1+Ch	0.4239	8	8	**0.7073**	0.0021	<.0001	**0.6606**	**0.8369**	**0.4389**	**0.4460**		0.0426
z_2A_2Ch	0.5120	9	9	**0.0826**	0.0003	<.0001	0.0363	**0.0929**	**0.5067**	**0.1274**	0.0426	

EAS = EP			**Least Squares means for effect EAS*HT Pr > |t| for H_0: LSMean(i) = LSMean(j) Dependent variable: WI (EAS = EP)**									
HT	**WI LSMEAN**	**i**	**i/j**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1A_0Ch	0.8923	1	1		**0.4188**	**0.3803**	0.0151	<.0001	<.0001	<.0001	<.0001	<.0001
1A_1+Ch	0.8738	2	2	0.4188		**0.1748**	**0.5274**	0.0412	<.0001	**0.0914**	0.0002	0.0165
2A(1+R)	0.9098	3	3	0.3803	0.1748		0.0063	<.0001	<.0001	<.0001	<.0001	<.0001
2A_0Ch	0.8600	4	4	0.0151	**0.5274**	0.0063		0.0032	<.0001	0.0092	<.0001	0.0003
2A_1Ch	0.8277	5	5	<.0001	0.0412	<.0001	0.0032		<.0001	**0.3757**	0.0007	**0.4833**
2A_3+Ch	0.7389	6	6	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	0.0037	<.0001
Other_0Ch	0.8371	7	7	<.0001	**0.0914**	<.0001	0.0092	**0.3757**	<.0001		<.0001	**0.1033**
Other_1+Ch	0.7913	8	8	<.0001	0.0002	<.0001	<.0001	0.0007	0.0037	<.0001		0.0071
z_2A_2Ch	0.8196	9	9	<.0001	0.0165	<.0001	0.0003	**0.4833**	<.0001	**0.1033**	0.0071	

Source: EU-SILC 2020, own processing in SAS EG.

of several types of households appeared. Therefore, to assess the impact of HT on the WI, a more comprehensive analysis for unemployed persons (UP) and employed persons (EP) was required, using a contrast analysis.

A comparison of pairs of marginal WI means for unemployed persons (UP) shows that there was no statistically significant difference ($p = 0.4108$) between the types of 1A_1$^+$Ch and 2A(1$^+$R) and these two types of households had a statistically significantly lower WI at the significance level of 0.05 as the other types of households. There was no significant difference between the other pairs of household types in terms of the WI at significance level of 0.01, so the study verified whether the WI means in these other household types (except 1A_1$^+$Ch and 2A(1$^+$R)) could be considered the same.

Denoting the WI means for the UP economic activity status by μ_{2i}, where index 2 expresses the 2$^{\text{nd}}$ category of the EAS factor (UP) and $i = 1, 2, \ldots, 9$ determines the type of household (see Table 1), then the subject of interest was the hypothesis:

$$H_0 : \mu_{21} = \mu_{24} = \mu_{25} = \mu_{26} = \mu_{27} = \mu_{28} = \mu_{29} .$$

To test it, a simultaneous test of six null hypotheses was used:

$$H_0 : \mu_{21} = \mu_{24} \ \wedge \ H_0 : \mu(\mu_{21}, \mu_{24}) = \mu_{25} \ \wedge \ H_0 : \mu(\mu_{21}, \mu_{24}, \mu_{25}) = \mu_{26} \ \wedge$$
$$H_0 : \mu(\mu_{21}, \mu_{24}, \mu_{25}, \mu_{26}) = \mu_{27} \ \wedge \ H_0 : \mu(\mu_{21}, \mu_{24}, \mu_{25}, \mu_{26}, \mu_{27}) = \mu_{28} \ \wedge$$
$$H_0 : \mu(\mu_{21}, \mu_{24}, \mu_{25}, \mu_{26}, \mu_{27}, \mu_{28}) = \mu_{29} ,$$

which was then rewritten into linear combinations

$$H_0 : \mu_{21} - \mu_{24} = 0 ,$$
$$H_0 : \tfrac{1}{2}\mu_{21} + \tfrac{1}{2}\mu_{24} - \mu_{25} = 0 ,$$
$$H_0 : \tfrac{1}{3}\mu_{21} + \tfrac{1}{3}\mu_{24} + \tfrac{1}{3}\mu_{25} - \mu_{26} = 0 ,$$
$$H_0 : \tfrac{1}{4}\mu_{21} + \tfrac{1}{4}\mu_{24} + \tfrac{1}{4}\mu_{25} + \tfrac{1}{4}\mu_{26} - \mu_{27} = 0 ,$$
$$H_0 : \tfrac{1}{5}\mu_{21} + \tfrac{1}{5}\mu_{24} + \tfrac{1}{5}\mu_{25} + \tfrac{1}{5}\mu_{26} + \tfrac{1}{5}\mu_{27} - \mu_{28} = 0 ,$$
$$H_0 : \tfrac{1}{6}\mu_{21} + \tfrac{1}{6}\mu_{24} + \tfrac{1}{6}\mu_{25} + \tfrac{1}{6}\mu_{26} + \tfrac{1}{6}\mu_{27} + \tfrac{1}{6}\mu_{28} - \mu_{29} = 0 .$$

The coefficients of linear combinations given in the null hypotheses were used to simultaneously test these hypotheses using the CONTRAST statement. The coefficients needed to test the last (sixth) partial hypothesis are shown in Table 4.

In the ‘total’ row in Table 4, there are coefficients for the HT factor, not used in the statement because the variable HT itself is not in the model (see Table 2). The coefficients for the EAS factor are zero, so the EAS factor itself did not enter the statement. The coefficients for interaction listed in the field of Table 4 were used. Similarly, one could determine the coefficients for the other five partial hypotheses. The coefficients determined in this way were used in the CONTRAST statement,

Table 4

Coefficients for the CONTRAST statement to test the null hypothesis
$H_0: \mu(\mu_{21}, \mu_{24}, \mu_{25}, \mu_{26}, \mu_{27}, \mu_{28}) = \mu_{29}$ for the $EAS \times HT$ interaction

EAS	***HT***									**Sum**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	
1	0	0	0	0	0	0	0	0	0	**0**
2	$\frac{1}{6}$	0	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	-1	**0**
3	0	0	0	0	0	0	0	0	0	**0**
Sum	$\frac{1}{6}$	**0**	**0**	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	**-1**	**0**

Source: own processing.

while the relevant variable and its associated coefficients for the individual partial hypotheses of simultaneous testing were separated by a comma:

```
CONTRAST '21=24=25=26=27=28=29'
         EAS*HT 0 0 0 0 0 0 0 0 0 1 0 0 -1,
         EAS*HT 0 0 0 0 0 0 0 0 0 0.5 0 0 0.5 -1,
         EAS*HT 0 0 0 0 0 0 0 0 0 0.3333 0 0 0.3333 0.3333 -1,
         EAS*HT 0 0 0 0 0 0 0 0 0 0.25 0 0 0.25 0.25 0.25 -1,
         EAS*HT 0 0 0 0 0 0 0 0 0 0.2 0 0 0.2 0.2 0.2 0.2 -1,
         EAS*HT 0 0 0 0 0 0 0 0 0 0.16666 0 0 0.16666 0.16666
                0.16666 0.16666 0.16666 -1;
```

Running this statement within PROC GLM generates the first line in Table 5.

Table 5

Simultaneous equality tests for LS-Means of WI for unemployed persons from selected types of households

Contrast	DF	Contrast SS	Mean square	F value	Pr > F
21=24=25=26=27=28=29	6	0.2496	0.0416	1.00	0.4200
22 vs 23	1	0.0280	0.0280	0.68	0.4108
21-24-25-26-27-28-29 vs 22-23	1	1.0583	1.0583	25.56	<.0001

Source: EU-SILC 2020, own processing in SAS EG.

Based on the result of a simultaneous test of six null hypotheses ($DF = 6$) at a significance level of 0.05, the null hypothesis ($p = 0.4200$) was not rejected, which means that in the case of unemployed persons there was not enough evidence to assume a different WI mean in households of type 1A_0Ch, 2A_0Ch, 2A_1Ch, 2A_2Ch, 2A_3$^+$Ch, Other_0Ch, and Other_1$^+$Ch. At the same time, there was also no significant difference ($p = 0.4108$) between the other two types of households:

1A_1+Ch and 2A(1+R), which were tested both with the CONTRAST statement (2nd row in Table 5) and the analysis of marginal means (Table 3). Thus, the study created two clusters of household types for unemployed persons, while there were no significant differences between the types of households belonging to the common cluster, but there were significant differences between the clusters ($p < 0.0001$; 3rd row in Table 5).

To estimate the WI mean in these two household clusters, the author used the ESTIMATE statement, illustrated in the estimate for the cluster of household types 1A_0Ch, 2A_0Ch, 2A_1Ch, 2A_2Ch, 2A_3+Ch, Other_0Ch, and Other_1+Ch. The coefficients for these types of households are shown in Table 6. As with the CONTRAST statement, the coefficients in the total row that apply to the HT variable were not used because this variable does not appear separately in the model. However, this time there were non-zero coefficients for the EAS variable (column sum) and a non-zero coefficient for the intercept (total sum of coefficients in the bottom right corner), which were written in the ESTIMATE statement. Since the intercept can only be counted once, the DIVISOR option with constant 7 was used.

Table 6

Coefficients for the ESTIMATE statement to estimate $\mu\left(\mu_{21},\mu_{24},\mu_{25},\mu_{26},\mu_{27},\mu_{28},\mu_{29}\right)$ for the $EAS\times HT$ interaction

EAS	***HT***									**Sum**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	
1	0	0	0	0	0	0	0	0	0	**0**
2	1	0	0	1	1	1	1	1	1	**7**
3	0	0	0	0	0	0	0	0	0	**0**
Sum	**1**	**0**	0	1	1	1	1	1	1	7

Source: own processing.

After running the statement

```
ESTIMATE '21-24-25-26-27-28-29' intercept 7 EAS 0 7 0
         EAS*HT 0 0 0 0 0 0 0 0 0 1 0 0 1 1 1 1 1 1 /divisor = 7;
```

within PROC GLM, one obtains the first row in Table 7.

Table 7

The estimate of $\mu\left(\mu_{21},\mu_{24},\mu_{25},\mu_{26},\mu_{27},\mu_{28},\mu_{29}\right)$ and $\mu\left(\mu_{22},\mu_{23}\right)$ for the $EAS\times HT$ interaction

Parameter	**Estimate**	**Standard error**	**t value**	**Pr > |t|**
21-24-25-26-27-28-29	0.4420	0.0247	17.87	<.0001
22-23	0.1003	0.0687	1.46	0.1443

Source: EU-SILC 2020, own processing in SAS EG.

In the case of an unemployed, there were no significant differences in terms of the WI mean person between household types 1A_0Ch, 2A_0Ch, 2A_1Ch, 2A_2Ch, 2A_3^{+}Ch, Other_0Ch, and Other_1^{+}Ch. In the cluster of these types of households, the author estimated the WI mean at 44.20%, and with a probability of 0.95 it was estimated in the interval (0.4420–1.9603×0.0247; 0.4420+1.9603×0.0247), i.e. (0.3936; 0.4904), using the quantile of the Student's distribution $t_{0.975}(7\,346) = 1.9603$. Similarly, the study estimated the WI mean for a cluster of household types 1A_1^{+}Ch and 2A(1^{+}R) (10.03%; 2nd row in Table 7). For unemployed persons, the WI mean across these two types of households was not significantly different from 0 ($p = 0.1443$) and with a risk of 0.075 did not exceed 20%, which is the limit determining a very low work intensity. In other words, a person who has the status of unemployed and lives in a household of type 1A_1^{+}Ch or 2A(1^{+}R), had in 2019 up to a 92.5% confidence level of showing a very low work intensity over the entire reference period.

If a person is employed, the riskiest types of households, in terms of exclusion from the labour market, are 2A_3^{+}Ch and Other_1^{+}Ch, between which there is a significant difference ($p = 0.0037$; to the detriment of 2A_3^{+}Ch). However, both types of households have a visibly lower mean of the WI than other types of households (Table 2; $p < 0.001$). This is followed by a cluster of household types 2A_1Ch, Other_0Ch, and 2A_2Ch, where there is no significant difference between the pairs of these types ($p = 0.3757$, $p = 0.4833$ and $p = 0.1033$) and based on the CONTRAST statement, we found out that there is also no significant difference between all 3 types of households ($p = 0.2577$; Table 8).

Table 8

Simultaneous test of equality of marginal WI means for employed persons from selected types of households

Contrast	DF	Contrast SS	Mean square	F value	Pr > F
35=37=39	2	0.1123	0.0561	1.36	0.2577

Source: EU-SILC 2020, own processing in SAS EG.

Furthermore, the combination of household types 1A_1^{+}Ch and 2A_0Ch ($p = 0.5274$) and household types 1A_0Ch and 2A(1^{+}R) ($p = 0.3803$) proved to be reasonable. For employed persons from the above five clusters of household types, the estimates of LS-means of the WI are presented in Table 9.

For all the five clusters of household types for employed persons, the WI mean was significantly different from 0 ($p < 0.0001$). In addition, there was a statistically significantly different WI mean between all pairs of these five clusters. The smallest difference was between persons in the household cluster 1A_1$^+$Ch, 2A_0Ch and persons in household cluster 1A_0Ch, 2A(1$^+$R). The author estimated this difference in the WI mean at 3.4%, but this was also significantly different from 0 at the significance level of 0.05, $p = 0.0196$.

Table 9

The estimate of μ_{36}, μ_{38}, $\mu\left(\mu_{35}, \mu_{37}, \mu_{39}\right)$, $\mu\left(\mu_{32}, \mu_{34}\right)$ and $\mu\left(\mu_{31}, \mu_{33}\right)$ for the *EAS*×*HT* interaction

Parameter	Estimate	Standard error	T value	Pr>|t|
2A_3$^+$Ch	0.7389	0.0191	38.75	<.0001
Other_1$^+$Ch	0.7913	0.0109	72.73	<.0001
2A_1Ch, Other_0Ch, 2A_2Ch	0.8281	0.0099	83.72	<.0001
1A_1$^+$Ch, 2A_0Ch	0.8669	0.0134	64.70	<.0001
1A_0Ch, 2A(1$^+$R)	0.9010	0.0135	66.80	<.0001

Source: EU-SILC 2020, own processing in SAS EG.

Following the above procedure, the study created five household clusters for other inactive persons (IP), two household clusters for unemployed persons (UP), five household clusters for employed persons (EP); these clusters are listed in Table 10.

Table 10

Household clusters for individual statuses of economic activity

EAS	Cluster	Household types
Inactive person	**IP 1**	1A_0Ch, 2A(1$^+$R)
	IP 2	2A_0Ch, Other_0Ch
	IP 3	Other_1$^+$Ch
	IP 4	2A_1Ch, 2A_3$^+$Ch
	IP 5	1A_1$^+$Ch, 2A_2Ch
Unemployed	**UP 1**	1A_1$^+$Ch, 2A_1Ch
	UP 2	1A_0Ch, 2A(1$^+$R), 2A_0Ch, 2A_2Ch, 2A_3$^+$Ch, Other_0Ch, Other_1$^+$Ch
Employed	**EP 1**	2A_3$^+$Ch
	EP 2	Other_1$^+$Ch
	EP 3	2A_1Ch, Other_0Ch, 2A_2Ch
	EP 4	1A_1$^+$Ch, 2A_0Ch
	EP 5	1A_0Ch, 2A(1$^+$R)

Source: own processing.

Point and interval (95%) estimates of the WI mean for persons from individual household clusters are shown in Figure 1.

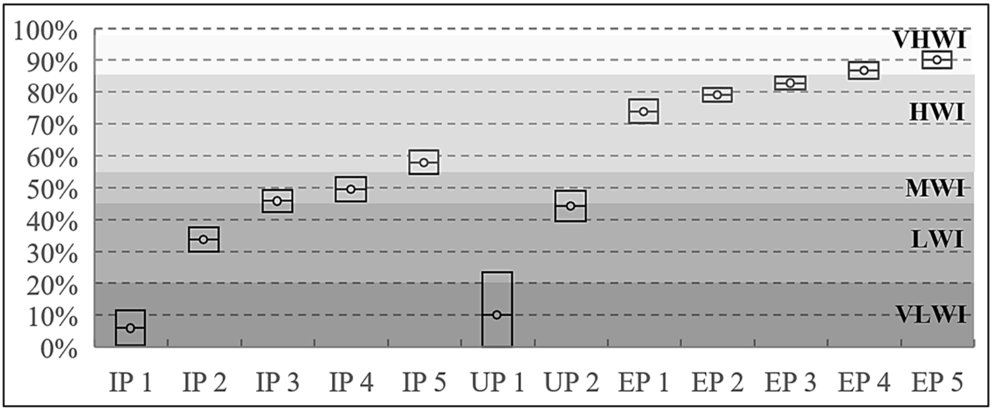

Fig. 1. Interval estimates (95%) of LS-means of WI for *EAS*×*HT* interaction

Source: EU-SILC 2020, own processing in SAS EG.

Households without dependent children (cluster IP 1 and IP 2) are the riskiest for other inactive persons in terms of exclusion from the labour market and within them, households with at most one person of working age (IP 1) are especially at risk. Inactive persons from households in cluster IP 1 (1A_0Ch, 2A(1$^+$R)) have the WI mean much below the upper limit for identifying a very low work intensity (below 20%). On the contrary, for inactive persons, the highest mean of WI was for persons from households of cluster IP 5 (1A_1$^+$Ch, 2A_2Ch), with the identified high work intensity. Other inactive persons from households belonging to other clusters (IP 2 to IP 4) showed the WI mean at the low or medium level.

The HT factor does not show consistent results for individual statuses of economic activity. While, e.g. inactive persons from households of type 1A_1$^+$Ch were not at risk in terms of exclusion from the labour market, the unemployed persons from this type of household were at high risk. Unemployed persons from households 1A_1$^+$Ch were put together with unemployed persons from households 2A_1Ch included in the common cluster UP 1, in which the WI mean was at the level of very low work intensity (with a low probability of reaching the level of low work intensity). The cluster of unemployed persons from other types of households (UP 2) was at the level of low to medium labour intensity. Unlike other inactive persons, unemployed persons in no type of household showed WI mean at a level higher than the medium, while employed persons had WI mean at a high to a very high level in all types of households. With 95% confidence, employed persons from cluster EP 5 (1A_0Ch, 2A(1$^+$R)), i.e. from households with no dependent children and at most one adult of working age, showed a WI mean at the very high level. It should be noted that persons from these types of households in the case of another inactive person, were in the riskiest position and their WI mean identified their (quasi)joblessness.

5.2.2. Analysis of the impact of interaction *EAS*×*Age*

As with the influence of the type of households, in the case of the influence of age on the WI, it was confirmed that employed persons have a significantly higher WI mean than unemployed persons or other inactive persons ($p < 0.0001$). There was a significant difference between unemployed persons and other inactive persons at the significance level of 0.05 only in the age group up to 30 years ($p = 0.0203$), to the detriment of unemployed persons. For other inactive persons, the lowest WI mean was in the age groups 30-40 and 40-50, between which there was no significant difference ($p = 0.9115$). Across these two categories (persons aged 30-50), the WI mean with 95% confidence level was in the range of 26.8-33.2% (IP 30-50 in Figure 2). Insignificant differences were also confirmed for unemployed persons in the age groups 30-40, 40-50, and 50-60 ($DF = 2$; $p = 0.6795$). Unemployed persons aged 30-60 years showed the WI mean with 95% confidence level from the interval 30.4-40.3% (UP 30-60 in Figure 2). There were significant differences between the other age categories for the individual statuses of economic activity in terms of the WI, and the point and interval estimates of the WI means are shown in Figure 2.

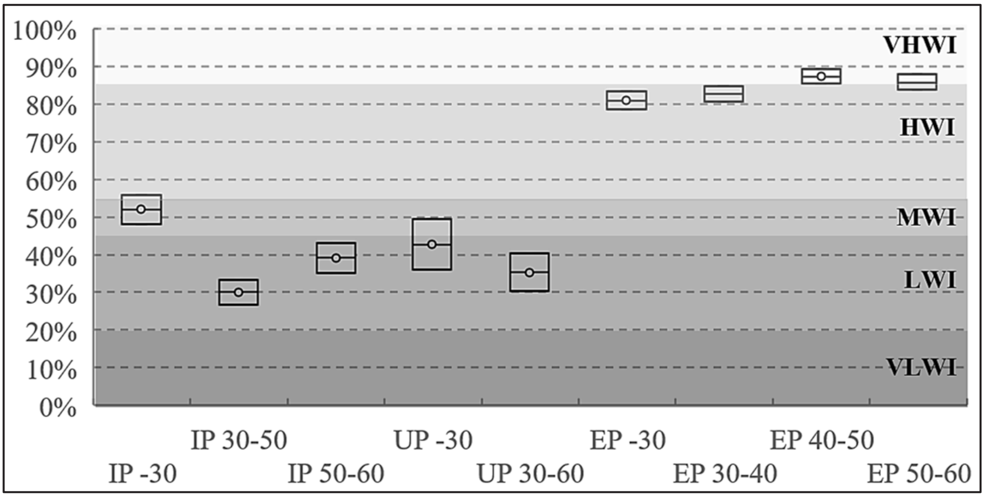

Fig. 2. Interval estimates (95%) of LS-Means of WI for *EAS*×*Age* interaction

Source: EU-SILC 2020, own processing in SAS EG.

Other inactive persons presented a demonstrably highest mean of the WI in the age group under 30 years, and this mean was at the medium level, while for older persons it was at the low level. The author observed a similar phenomenon for the unemployed, however for persons under the age of 30 one cannot convincingly assume the WI mean at the medium level, but at the low to medium level (Figure 2). While the age group of under 30 was the least at risk for other inactive people and unemployed people, this category (EP-30) showed the highest risk for employed people. Employed persons under the age of 30 may have the WI mean below 80% with a risk of 0.05, which was not the case for older employed persons. Despite this finding, employed persons under the age of 30 and also those aged 30-40, had the WI mean at the same level, specifically the high level. Employed persons aged 50-60 showed the WI mean at the level of high to very high, and employed persons aged 40-50 even a very high level (with a reliability of 0.95 from the interval 85.3-89.3%).

5.2.3. Analysis of the impact of interaction *EAS*×*Education*

Not surprisingly, higher WI is associated with higher education, and this applies to all statuses of economic activity (Figure 3). The differences in the WI mean between the ISCED 3-5 and ISCED 6-8 educational groups were not as large and convincing ($p=0.0084$ for IP, $p=0.0329$ for UP, $p=0.0386$ for EP as the differences between these two educational groups from ISCED 0-2 ($p < 0.0001$). For other inactive persons (IP), there were maller differences between the educational groups in the

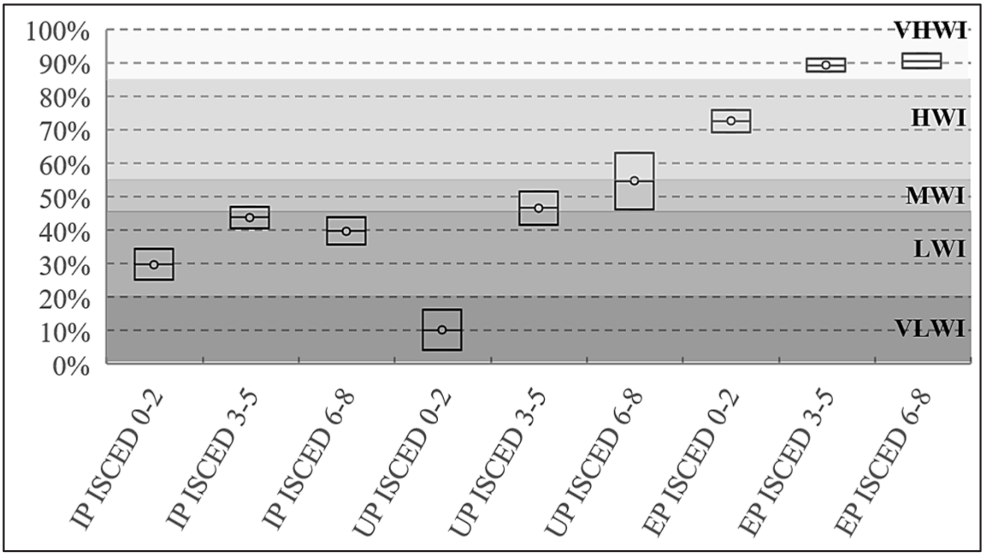

Fig. 3. Interval estimates (95%) of LS-Means of WI for *EAS*×*Education* interaction

Source: EU-SILC 2020, own processing in SAS EG.

WI mean than in the other two statuses. Although other inactive persons with education at ISCED 0-2 level showed a significantly lower WI mean than persons with higher education, all the educational groups for this status of economic activity had the WI mean at the low level. The largest exclusion from the labour market naturally concerns the unemployed with low education (ISCED 0-2), for whom the WI mean was estimated at a very low level.

As unemployed persons with higher education showed the WI mean at a low to medium (ISCED 3-5) or medium to a high level (ISCED 6-8), they seem to remain unemployed for a shorter time and/or have another adult in a household (spouse or partner) who is employed. However, the status of economic activity – Other inactive or Unemployed – had such a significant effect on the WI that education cannot sufficiently compensate for this and therefore persons who are unemployed or other inactive and whose education is tertiary generally live in households with significantly lower WI ($p < 0.0001$) than employed persons. This applies even to employed persons with ISCED 0-2 education, for whom the average WI was at a high level. Employed persons with higher education have the WI mean even at a very high level.

Conclusion

The type of household in which a person lives, education, age, health conditions, marital status, as well as the region in which a person lives, are factors that significantly affect the WI, but this impact varies for different statuses of economic activity. These factors are considered very important by other researchers (Horemans, 2018; Verbunt and Guio, 2019) when assessing employment/unemployment and atypical employment in relation to social exclusion.

The general linear model, in which the study considered the interactions of these relevant factors with the status of economic activity of the assessed persons, explained the variability of the WI to more than 50%. The status of economic activity, type of household, age, and education participated the most in this explained variability (the participation of all the four factors was 48% in total). In addition to economic activity, van der Zwan and de Beer (2021) also used the above factors (age, education, and type of household) as control variables, applying them to the assessment of employment for persons with disability and those without disability. Horemans (2016) discussed the importance of the impact of age and education on non-standard employment, while also describing the need to transform these factors at the individual level to the household level, through the degree of educational and age homogamy.

The paper focuses on the analysis of the influence of the above four factors (status of economic activity, type of household, age, and education) on the WI, to which the

LS-means and contrast analysis within the estimated GLM was used. Through the equality tests of LS-means of WI and simultaneous testing of WI means, the author identified for each economic activity status between which types of households, age categories, and educational categories, there was no significant difference from the WI perspective, and between which there were demonstrable differences. For household types, age groups, and educational groups for which the author did not have sufficient evidence to be able to demonstrably confirm differences in the mean of WI, such clusters were created between which there were statistically significant differences. Individually for employed persons, unemployed persons, and other inactive persons, the mean of WI was estimated for individual types of households, age groups, and educational categories (or their groups). Based on interval estimates of the WI mean, these groups of persons were assigned a level of WI in accordance with the Eurostat methodology.

For other inactive people, the study mainly identified very low and low levels of WI. The WI mean for other inactive persons can be assumed to be above 50% only exceptionally, and this applies to persons under the age of 30, persons from complete households with dependent children, and persons from households with 1 adult and at least 1 child. The study identified very low work intensity ((quasi)joblessness) most often for unemployed persons. For the unemployed, the author estimated the work intensity to be below 50% for all types of households and all age groups. Unemployed people are highly likely to live in households that use more than 50% of their employment potential only if they have a tertiary education. For employed persons, broken down separately by type of household, age, and education, the study revealed the WI mean at least at a high level.

In terms of household type, other inactive persons from households without dependent children, in which there is at most one person of working age, and unemployed persons from households of 1 adult with at least 1 dependent child and households of 2 adults with 1 child, have the greatest risk of exclusion from the labour market. For these groups of persons, the study quantified the WI mean at a very low level. Persons from households without dependent children, in which there is at most one person of working age, were most at risk if they had the status of economic activity other than inactive, but in the case of the status of employed they reached the highest WI mean, which was at a very high level.

The conducted analyses revealed that of all age groups, the under-30 age group had the highest WI mean when considering unemployed persons (UP) or other inactive persons (IP), for whom the mean WI was at a low to medium level (for UP) and at a medium level (for IP). Older unemployed and other inactive persons generally showed low work intensity. Employed persons had the WI mean significantly higher than the other two statuses of economic activity, and this was also confirmed in the case of a breakdown of persons according to their age.

Depending on their age, employed persons had a high to a very high WI mean (up to 40 years – high, 40-50 years – very high, and 50-60 years – high to very high).

With the increase in education, the WI was also growing. Unemployed persons with low education (ISCED 0-2) generally showed (quasi)joblessness, while unemployed people with tertiary education were 50% more likely to have a medium WI, and 50% to even have a high WI. In this way, the study revealed that education plays a crucial role for unemployed persons in terms of work intensity. These results are consistent with the assumption that persons with higher education are less likely to remain unemployed for a longer period (Núñez and Livanos, 2010) and that they usually have a person with higher education (educational homogamy) at their side, whose threat of exclusion from the labour market is lower compared to a less educated person. Education for other inactive persons does not cause such large differences in WI as for the unemployed. Other inactive persons in all educational groups fall into the low level according to the WI mean. Employed persons have the WI mean in terms of education at a high level (for ISCED education 0-2) or a very high level (for ISCED education 3-8).

The author is convinced that the paper fills a gap in research that mostly focuses only on labour market exclusion identified on the basis of very low work intensity. The comprehensive work intensity analysis provided in this paper is important because also other degrees of work intensity may be associated with poverty and social exclusion (Kis and Gábos, 2016, Kalinowski, 2018). Social policies should then be targeted at persons with reduced work intensity, as confirmed by Blatná (2018). Based on the analysis of the share of people living in households with very low work intensity in the Czech Republic in the period 2005-2016, that study found that the growth in social benefits and increase in the proportion of people in lifelong education leads to a reduction of the proportion of people living in (quasi)jobless households.

In conclusion, it should be emphasised that the paper provides an empirical analysis for Slovakia, and although many conclusions apply at least to the CEE countries, this needs to be validated by further research. In particular, the influence of household type on the WI can be significantly different in other countries. According to Atkinson et al. (2017), there is a great deal of cross-country variation in the composition of the (quasi)jobless population by household type. The results of the analysis have their limitations, mainly related to the methodology of measuring work intensity and its levels (specifically very low work intensity). Ward and Özdemir (2013) noted a problem in the definition of work potential (the denominator of the WI indicator), which does not include persons older than 59, as well as in the threshold for identifying (quasi)jobless households that is set at 20%, whereas they advocate its raising to 30%.

References

Agresti, A. (2015). Foundations of Linear and Generalized Linear Models. New York: John Wiley & Sons.

Atkinson, T., Guio, A. C., and Marlier, E. (eds.) (2017). Monitoring social inclusion in Europe. *Statistical Books Eurostat*. Luxembourg: Publications Office of the European Union.

Blatná, D. (2018). Analysis of the relationship between the share of people living in households with very low work intensity and selected socio-economic indicators in the Czech Republic in the period 2005-2016. 12th International Days of Statistics and Economics, Prague.

Cai. W. (2014). Making Comparisons Fair: How LS-Means Unify the Analysis of Linear Models. *SAS Institute Inc. Paper SA*, S060–2014.

Calegari, E., Fabrizi, E., and Mussida, C. (2021). Disability and work intensity in Italian households. *Review of Economics of the Household*, 1–20.

Cantó, O., Cebrián, I., and Moreno, G. (2022). Youth living arrangements and household employment deprivation: Evidence from Spain. *Journal of Family Research. Early View*, 1–33.

Colombarolli, C. (2021). In-work poverty and regional disparities. An analysis of the relationship between work intensity and the probability of being and feeling poor across Italian territories. *Sociologia del lavoro*, 161, 76–93.

Ćwiek, M. T., and Ulman, P. (2019). Income and poverty in households in selected European countries. *Acta Universitatis Lodziensis. Folia Oeconomica*, 345(6), 9–34.

Darlington, R. B., and Hayes, A. F. (2016). Regression Analysis and Linear Models: Concepts, Applications, and Implementation. Guilford Publications.

de Graaf-Zijl, M., and Nolan, B. (2011). Household joblessness and its impact on poverty and deprivation in Europe. *Journal of European Social Policy*, 21(5), 413–431.

Dean, A., Voss, D., and Draguljić, D. (2017). Design and Analysis of Experiments. Springer.

Duiella, M., and Turrini, A. (2014). Poverty developments in the EU after the crisis: A look at main drivers. *ECFIN Economic Brief*, 31, 1–10.

Elswick Jr, R. K., Gennings, C., Chinchilli, V. M., and Dawson, K. S. (1991). A simple approach for finding estimable functions in linear models. *The American Statistician*, 45(1), 51–53.

Eurostat, Glossary: At risk of poverty or social exclusion (AROPE), https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:At_risk_of_poverty_or_social_exclusion_(AROPE), accessed on 6.2.2022, 2022.

Fabrizi, E., and Mussida, C. (2020). Assessing poverty persistence in households with children. *The Journal of Economic Inequality*, 18(4), 551–569.

Filandri, M., Pasqua, S., and Struffolino, E. (2020). Being working poor or feeling working poor? The role of work intensity and job stability for subjective poverty. *Social Indicators Research*, 147(3), 781-803.

García-Gómez, C., Pérez, A., and Prieto-Alaiz, M. (2021). Copula-based analysis of multivariate dependence patterns between dimensions of poverty in Europe. *Review of Income and Wealth*, 67(1), 165–195.

Gerbery, D., and Miklošovic, T. (2020). Labour market transitions and their determinants in Slovakia: Path from crisis to recovery 1. *Ekonomický časopis*, 68(7), 651–676.

Goodnight, J. H., and Harvey, W. R. (1997). SAS Technical Report R-103. Least Squares Means in The Fixed Effects General Model. Cary, NC: SAS Institute Inc.

Guio, A.-C., Vandenbroucke, F. (2019). Poverty and Child Deprivation in Belgium. A Comparison of Risk Factors in the Three Regions and Neighbouring Countries, A Comparison of Risk Factors in the Three Regions and Neighbouring Countries (March 1, 2019). King Baudouin Foundation.

Hick, R., and Lanau, A. In-work poverty in the UK: Problem, policy analysis and platform for action, Final Report, Cardiff University, https://orca.cf.ac.uk/103013/1/Hick%20and%20Lanau%20_%20InWork%20Poverty%20in%20the%20UK.pdf, accessed on 5.3.2022, 2017.

Horemans, J. (2016). Polarisation of non-standard employment in Europe: Exploring a missing piece of the inequality puzzle. *Social Indicators Research*, 125(1), 171–189.

Horemans, J. (2018). Atypical employment and in-work poverty. In *Handbook on In-Work Poverty*. Edward Elgar Publishing.

Kalinowski, S. (2018). The Working Poor in the European Union. Conference: International Scientific Days 2018, „Towards Productive, Sustainable and Resilient Global Agriculture and Food Systems".

Kim, K., and Timm, N. (2006). Univariate and Multivariate General Linear Models: Theory and Applications with SAS. Chapman and Hall/CRC.

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

Kuznetsova. A., Brockhoff. P. B., and Christensen. R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13).

Lenth, R., V. (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software*, 69(1), 1–33.

Littell, R. C., Stroup, W. W., and Freund, R. J. (2010). SAS for Linear Models. 4th ed. Cary, NC: SAS Institute Inc.

McFarquhar, M. (2016). Testable hypotheses for unbalanced neuroimaging data. *Frontiers in Neuroscience*, 10, 270.

Mysíková, M., Želinský, T., Garner, T. I., and Večerník, J. (2019). Subjective perceptions of poverty and objective economic conditions: Czechia and Slovakia a quarter century after the dissolution of Czechoslovakia. *Social Indicators Research*, 145(2), 523–550.

Nieuwenhuis, R., and Maldonado, L. (2018). The triple bind of single-parent families: Resources, employment and policies to improve well-being. *Policy Press*.

Núñez, I., and Livanos, I. (2010). Higher education and unemployment in Europe: An analysis of the academic subject and national effects. *Higher Education*, 59(4), 475–487.

Peña-Casas, R., Ghailani, D., Spasova, S., and Vanhercke, B. (2019). In-work poverty in Europe. A study of national policies. European Social Policy Network. Brussels: European Commission.

Poline, J. B., Kherif, F., Pallier, C., and Penny, W. (2007). Contrasts and classical inference. *Statistical Parametric Mapping: the Analysis of Functional Brain Images*, 1, 126–139.

SAS Institute Inc., SAS/STAT® 15.1 User's Guide. The GENMOD Procedure, Cary SAS/STAT® 15.1, NC: SAS Institute Inc., 2018a.

SAS Institute Inc., SAS/STAT® 15.1 User's Guide. The GLM Procedure. Cary SAS/STAT® 15.1, NC: SAS Institute Inc., 2018b.

Sánchez-Sellero, M. C., and Garcia-Carro, B. (2020). Which groups have a greater risk of poverty in Spain? *Revija za socijalnu politiku*, 27(1), 36–36.

Schad, D. J., Vasishth, S., Hohenstein, S., and Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038.

Searle, S. R., and Gruber, M. H. J. (2017). Linear Models, 2nd ed. John Wiley & Sons.

Searle, S. R., Speed, F. M., and Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician*, 34(4), 216–221.

Suzuki, M., Taniguchi, T., Furihata, R., Yoshita, K., Arai, Y., Yoshiike, N., and Uchiyama, M. (2019). Seasonal changes in sleep duration and sleep problems: A prospective study in Japanese community residents. *PLoS One*, 14(4), e0215345.

Šoltés, E., and Ulman, P. (2015). Material deprivation in Poland and Slovakia – A comparative analysis. *Cracow Review of Economics and Management*, 11(947), 19–36.

Tabachnick, B. G., and Fidell, L. S. (2013). Using Multivariate Statistics, 6th ed. Boston, MA: Pearson.

van der Zwan, R., and de Beer, P. (2021). The disability employment gap in European countries: What is the role of labour market policy? *Journal of European Social Policy*, 31(4), 473–486.

Verbunt, P., and Guio, A.-C. (2019). Explaining differences within and between countries in the risk of income poverty and severe material deprivation: Comparing single and multilevel analyses. *Social Indicators Research*, 144(2), 827–868.

Wang, B., Wu, P., Kwan, B., Tu, M. X., and Feng, Ch. (2018). Simpson's paradox: Examples. *Shanghai Archives of Psychiatry*, 30(2), 139.

Ward, T., and Ozdemir, E. (2013). Measuring low work intensity – An analysis of the indicator. *ImPRovE Discussion Paper*, 13/09, Antwerp.

Westfall, P. H., and Tobias, R. D. (2007). Multiple testing of general contrasts: Truncated closure and the extended Shaffer-Royen method. *Journal of the American Statistical Association*, 102(478), 487–494.

Received: June 2022

Acknowledgments: *This work was supported by the VEGA projects The impact of the COVID-19 crisis on business demography and employment in the Slovak Republic and the EU (No. 1/0561/21) and Poverty and Social Exclusion in Slovakia and the EU in Times of the COVID-19 Pandemic and Energy Crisis (No. 1/0038/23).*