Comparison of (Quasi-)Joblessness in Slovakia and the Czech Republic through the Marginal Means based on Logit Models

Martina Košíková[1] | *University of Economics in Bratislava, Bratislava, Slovakia*
Erik Šoltés[2] | *University of Economics in Bratislava, Bratislava, Slovakia*

Received 26.2.2023, **Accepted** (reviewed) 19.4.2023, **Published** 15.12.2023

Abstract

In the paper, we analyze one of the aspects of measuring poverty and social exclusion in the world within the context of the Europe 2030 Strategy, which is represented by very low work intensity. Using the data obtained from the EU-SILC 2021 statistical survey (for Slovakia and the Czech Republic), we apply logistic regression methods and generalized linear models to quantify the impact of relevant categorical factors on the binary dependent variable very low work intensity of Slovak and Czech households. Based on the obtained results, we then process a comparative analysis, through which we quantify the same and also different features of these countries in terms of (quasi-)joblessness.

Keywords	***DOI***	***JEL code***
(Quasi-)joblessness, logistic regression, least squares means, work intensity	*https://doi.org/10.54694/stat.2023.8*	*C12, C21, E24*

INTRODUCTION

Currently, poverty and social exclusion are two of the global problems affecting the quality of life of people around world. Poverty and social exclusion are generally considered a socio-economic problem affecting an individual or a household, and these social phenomena occur to a greater or lesser extent in every country in the world. Šoltés et al. (2018) define that poverty and social exclusion affect people's quality of life and are assessed from several aspects, such as the amount of equivalent disposable income (income poverty), material deprivation or very low work intensity, which is often referred to as (quasi-) joblessness (QJ).[3]

[1] Department of Statistics, Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia. E-mail: martina.kosikova@euba.sk.
[2] Department of Statistics, Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia. E-mail: erik.soltes@euba.sk. ORCID: <https://orcid.org/0000-0001-8570-6536>.
[3] In the following text, we will also use the abbreviation QJ to denote (quasi-)joblessness.

In the paper, we focus on the risk that a person will live in a household with very low intensity, which is one of the three dimensions of poverty and social exclusion based on the concept of the EU 2030 Strategy. Work intensity is defined as the rate at which work potential is used in the household. The Eurostat methodology (2021) defines that work intensity represents the ratio of the number of months worked by individual household members of productive age (18–64 years) to the total number of months (within the reference period) that household members could have worked (persons of productive age don‘t include students aged 18 to 24). The work intensity indicator takes on values from 0, respectively 0% (none of the household members of productive age worked in the reference period) after 1, or 100% (every household member of productive age worked during the entire reference period). Households in which the work potential was used to less than 20% are considered as households with very low work intensity. Other levels of work intensity are discussed, for example by Eurostat (2022a), or Šoltés and Vojtková (2018). Households with very low work intensity are also referred to as (quasi-) joblessness households. According to Vlačuha and Kováčová (2021), (quasi-)joblesness households have a demonstrably higher risk of poverty than households with high or very high work intensity. This dimension of poverty and social exclusion should be particularly monitored because it significantly affects the poverty and social exclusion of children, which was also confirmed by the authors De Graaf-Zijl and Nolan (2011), who defined in their studies that (quasi-)joblessness in most countries is closely associated with child deprivation.

Many studies have reported that the risk of poverty and social exclusion depends on various measurable factors; for example, according to Duiella and Turrini (2014) these are factors such as gender, age, migration, geographical region. Eurostat (2022b) analyzed the risk of poverty or social exclusion based on socio-economic statistics such as not only gender and age, but also education and economic activity. The authors Johnston and McGauran (2018) in their studies analyzed (quasi-)unemployment depending on the type of households, economic activity and other socio-economic factors (level of education, age, gender, disability, etc.).

The results of Eurostat (2022b) showed that approximately 15.6% of the population in Slovakia was at risk of poverty in 2021 and Slovakia along with the Czech Republic (10.7%) are among the EU countries with the lowest poverty risk rate. However, there is a high risk of poverty in (quasi-)joblessness households in Slovakia (76.5%) and the Czech Republic (58.6%).

In the paper, we analyze the risk that a person will live in a (quasi-)joblessness household in Slovakia and the Czech Republic. The goal is to assess the impact of relevant factors on this risk in both countries individually and at the same time, to compare the results of the analyses. We mainly focus on:

1. assessment of the statistical significance of the impact of selected factors on the probability of living in a (quasi-)joblessness household,
2. identify the categories of which factor between which there are no significant differences and those categories or groups of categories between which there are demonstrable differences in terms of the risk of (quasi-)joblessness,
3. identify the risk g roups of persons in terms of the risk of (quasi-)joblessness,
4. comparison ofprobability of (quasi-)joblessness for different profiles of people determined by the three most fundamental factors.

All the stated goals are implemented specifically for persons living in Slovak households and specifically for persons living in Czech households. The paper further provides a comparison between Slovakia and the Czech Republic according to the mentioned attributes.

1 LITERATURE REVIEW

Poverty is a complex, multidimensional phenomenon that cannot be expressed only on the basis of one aspect or one definition. Poverty is primarily a socio-economic problem that has affected human society since its inception. There are several concepts of poverty that are dealt with in various scientific works. For example, the authors Atkinson et al. (2017), who analyzed poverty indicators based on the EU-SILC survey with macroeconomic analysis of aggregates within the EU countries and found that the EU did not achieve any positive change or progress in achieving the goal in the Europe 2020 Strategy (until 2020 the number of people at risk of poverty and social exclusion was reduced at least by 20 million). Ravallion (2020) stated in his study that over the last 30 years, poverty has decreased globally, which is mainly due to the reduction of absolute poverty in the developing world. Based on Brown and Long (2018), absolute poverty refers to people who lack their basic physical needs. Nolan and Marx (2009) assessed the factors affecting poverty differences in the OECD countries. Lafuente et al. (2020) evaluated the convergence of poverty between the EU countries from the perspective of the three dimensions of poverty and social exclusion in the concept of the Europe 2020 Strategy, finding that convergence in the EU countries exists and takes place based on a long-term process, in which the levels of individual member countries with similar characteristics are gradually approaching the levels of more developed countries.

OECD (2009) confirmed the existence of a close relationship between the level of personal income and the intensity of household work. On the contrary, De Graaf-Zijl and Nolan (2011) stated in their study that joblessness is not directly related to income poverty and pointed out that the risk of (quasi-) joblessness most often occurs among persons with a lower education, a disability or in a single-person household.

The selection of factors affecting (quasi-)joblessness was determined by the results of our previous research (e.g. Šoltés et al., 2022) and studies that analyzed this issue, such as Guio et al. (2022), who dealt with the analysis of the determinants of child deprivation in 31 European countries, or Johnston and McGauran (2018), who assessed the impact of relevant socio-economic factors on very low work intensity. Duiella and Turrini (2014) stated that for a more accurate analysis of (quasi-)joblessness it was important to consider socio-demographic and economic indicators such as age, gender, type of household or education. Since (quasi-) joblessness is closely related to other dimensions of poverty and social exclusion, when choosing potential factors we were also inspired by scientific works that assessed income poverty and material deprivation. Filandri and Strufolino (2019) found that between the indicators affecting the risk of working poverty were a low degree of education and the young age of persons, while households with a larger number of children and a small number of economically active persons are also at risk.

In the paper, we analyze (quasi-)joblessness using logistic regression. Logistic regression is quite popular in the study of poverty and social exclusion, which is confirmed by a number of studies. For example, Dudek and Lisicka (2013) in their study analyzed the incidence of poverty risk through binary logistic regression. Using logistic regression, Rusnak (2012) identified indicators that increase poverty. Saccone and Deaglio (2020) compared poor and developing economies with high-income economies using logit analysis. Gallardo (2020) analyzed multidimensional poverty using probit and logit models. The adequacy of applying logit models in the area of social exclusion was also confirmed by Gao et al. (2022), who used these models to identify poor households, than Stanley et al. (2011) who analyzed the factors that significantly contribute to the increased risk of social exclusion or Bradshaw et al. (2000) who assessed the relationship between poverty and social exclusion through logit models.

As we stated, logistic regression models are relatively extended in poverty and social exclusion research, but their potential is mostly not used to the full extent. A deeper analysis of the impact of individual factors on which dimension of poverty and social exclusion is provided by a contrast analysis, which is bound to such a model. For this reason, in our research on (quasi-)joblessness, we use not only the logit model, but also the contrast analysis, from which the most important conclusions presented in the article stem.

2 METHODS

Logistic regression falls into a broad class of generalized linear models. In logistic regression, the target variable has a binomial distribution, the systematic component is in the form of a linear combination of the explanatory variables, and the linear function is the logit.

The logistic regression model quantifies the log-odds for the "1" level of the binomial dependent variable Y as a function of the explanatory variables. Odds represents the probability ratio p_i, that the observed event occurs ($Y = 1$) against the probability $1 - p_i$, that observed event does not occur ($Y = 0$). Odds is expressed by the relation (1) and the log-odds, also referred to as logit, is given by the relation (2):

$$odds = \frac{p_i}{1 - p_i}, \tag{1}$$

$$logit\,(p_i) = ln\left(\frac{p_i}{1 - p_i}\right). \tag{2}$$

After transforming the qualitative variable Y (0 or 1) into a continuous variable expressed by the logit, the relation between the explanatory variable and the vector of explanatory variables is linear. So, the binary logistic regression model has the form:

$$logit\,(p_i) = ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} \dots + \beta_k x_{ik}, \tag{3}$$

where β_j are the unknown model parameters estimated using an iterative maximum likelihood method and x_{ij} are the observed values of the explanatory variables (where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, k$).

The significance of the logistic regression model is verified using three chi-square tests (likelihood ratio test, score test or Wald test), while Allison (2012) states that in the case of sufficiently large datasets it is not necessary to prefer any of the three tests. To verify the significance of the impact of independent variables on the dependent variable, the Wald test is used:

$$Wald = \hat{\boldsymbol{\beta}}^{\mathrm{T}} \cdot \mathbf{S}_{\mathrm{b}}^{-1} \cdot \hat{\boldsymbol{\beta}}, \tag{4}$$

where $\hat{\boldsymbol{\beta}}$ is the vector of the estimated regression coefficient and $\mathbf{S}_{\mathrm{b}}^{-1}$ is the variance-covariance matrix of the vector $\hat{\boldsymbol{\beta}}$. The Wald statistics has an asymptotically chi-squared distribution. Whereas, in the case of a quantitative variable, the number of degrees of freedom of this chi-square distribution is equal to 1 and in the case of a categorical factor, the number of degrees of freedom is equal to (r – 1), where *r* is the number of categories of the factor whose impact we are verifying.

An advantage of using logistic regression is that regression coefficients can be interpreted through odds ratios. The odds ratio is estimated from the logit model using an exponential transformation of the estimated regression coefficient $\hat{\beta}_j$, as follows

$$\hat{OR}_i = \frac{odds_1}{odds_2} = e^{\hat{\beta}_j}. \tag{5}$$

The odds ratio belonging to an independent continuous numerical variable expresses the extent to which the odd of the observed event changes with a unit increase in this numerical variable, under ceteris paribus conditions. In the case of categorical explanatory variables (categorical factors), the interpretation of the estimated odds ratios depends on the type of coding of the factors. In the case of referential coding, we assess the impact of each non-referential category on odd relative to the reference category. Effect coding compares the effects of individual categories to the mean impact across all categories of a factor.

By analyzing marginal means, we can identify hidden relationships between categories of independent variables in generalized linear models. If the estimated model includes qualitative independent variables that have more than two categories, analysis of marginal means allows comparison of differences in marginal means between pairs of categories. Marginal means, also referred to as LS means (least squares means) or EM means (estimated marginal means), are generally preferred over conditional means because they correct for data unbalance. LS means disregard the frequency of individual levels of the factor and when quantifying the means assign equal weight to each category, while arithmetic mean assigns individual weight to each category of the factor, which is conditioned by its frequency. Based on Suzuki et al. (2019) we note that while arithmetic means are estimated from the data, marginal means are estimated directly from the model. Wang et al. (2018) stated that for unbalanced datasets with more categorical or numerical factors, arithmetic means do not provide an adequate picture of the effect of the dependent variable for a particular factor because they ignore other effects, which may subsequently lead to Simpson's paradox. This paradox points to the importance of not ignoring causal relationships between variables (cause and effect) or confounding effects leading to incorrect conclusions in statistical analyses.

However, if we want to assess the significance of the differences in the marginal mean values of the independent variable between more than two categories at the same time, we use contrast analysis for this purpose. The basis of contrast analysis is general linear hypothesis (GLH) testing. We test hypotheses in the form of general linear hypotheses

$$H_0: c_1\mu_1 + c_2\mu_2 + \dots + c_p\mu_p = 0\,, \tag{6}$$

where μ_j is the marginal mean value of the independent variable for the jth level of the factor whose impact we are assessing (Batzer, 2020). These hypotheses are tested through a linear combination $c_1\beta_1 + c_2\beta_2 + \dots + c_p\beta_p$ parameters of the model β_j.

To test GLH, logistic regression uses the Wald statistic, which has an asymptotically chi-square distribution, which has one degree of freedom in the case of testing one linear combination and in the case of a simultaneous test of several linear combinations, the number of degrees of freedom is l, where l is the number of partial hypotheses that enter to the simultaneous test. Littell et al. (2010) provide additional information.

The LOGISTIC and GENMOD procedures in SAS Enterprise Guide statistical software are used in the paper to estimate logit models. The analysis of marginal means is realized in the SAS programming language through the LSMEANS statement in PROC GENMOD and contrast analysis was applied using the CONTRAST statement in PROC LOGISTIC. To identify the risk profiles of people in terms of (quasi-) joblessness in Slovakia and the Czech Republic and for the subsequent comparison between countries, the probabilities of living in a (quasi-)joblessness household were estimated in the paper. Point and interval estimates of the subject probabilities were quantified using the ESTIMATE statement. In more detail about the procedure of adequate testing of marginal means and their estimation by CONTRAST and ESTIMATE statement, e.g. Littell et al. (2010) a SAS Institute Inc. (2018).

3 DATABASE

The analyses presented in the paper are based on the EU-SILC 2021 survey. The input databases were provided by the statistical offices of Slovakia and the Czech Republic and consist of data relating to persons in Slovak and Czech households. The target variable is very low work intensity (VLWI). It is a binary variable with levels of "no" for persons not living in a household with very low work intensity and "yes" for persons living in a household with very low work intensity. The explanatory variables included in the model are listed in Table 1.

Table 1 Information about input factors and their categories

Factor	Levels	Description
EA (Economic Activity)	*Disabled_person* *Inactive_person* *Person_in_household* *Student* *Unemployed* *z_at_Work*	Disabled person, Person unable to work Other inactive person Person in the household Student Unemployed Employed
EDUCATION	*ISCED 0–2* *ISCED 3* *ISCED 4–5* *ISCED 6* *ISCED 7–8*	Primary and Lower secondary education Upper secondary education Post-secondary and Short-cycle tertiary education Bachelor education Master's or Doctorate education
HT (Type of Household)	*1A_0Ch* *1A_1+Ch* *2A(1+R)* *2A_0Ch* *2A_1Ch* *2A_3+Ch* *Other_0Ch* *Other_1+Ch* *z_2A_2Ch*	Household of 1 adult without dependent children Household of 1 adult with at least 1 dependent child Household of 2 adults with at least 1 aged 65+ Household of 2 adults without dependent children Household of 2 adults with 1 dependent child Household of 2 adults with at least 3 dependent children Other household without dependent children Other household with dependent children Household of 2 adults with 2 dependent children
MARITAL STATUS	*Divorced* *Never_married* *Widowed* *z_Married*	Divorced person Single person Widowed person Married person
AGE	*30–40* *40–50* *50+* *z_–30*	People aged between 31 and 40 years People aged between 41 and 50 years People aged 51 and over People under the age of 30
URB (Urbanisation)	*Sparse* *Intermediate* *z_Dense*	Sparsely populated area Intermediate area Densely populated area

Source: EU-SILC 2021 SO SR and CR, own processing

4 RESULTS

4.1 Analysis of the significance of the input factors impact on (quasi-)joblessness in Slovakia and the Czech Republic and modification of input factors

In addition to the factors listed in Section 3, explanatory factors such as gender, region and health status were also included in the analysis, but the method of stepwise elimination showed that these factors did

Table 2 Verification of the factors impact on (quasi-)joblessness

Type 3 Analysis of effects

Effect	DF	Slovakia		Czech Republic	
		Wald Chi-Square	Pr > ChiSq	Wald Chi-Square	Pr > ChiSq
EA	5	548.3650	<.0001	709.0464	<.0001
Education	4	116.9450	<.0001	12.2144	0.0158
HT	8	121.7755	<.0001	172.5083	<.0001
Marital_status	3	6.5516	0.0876	15.7331	0.0013
Age	3	8.2725	0.0407	29.0038	<.0001
Urbanisation	2	28.5390	<.0001	7.9296	0.0190

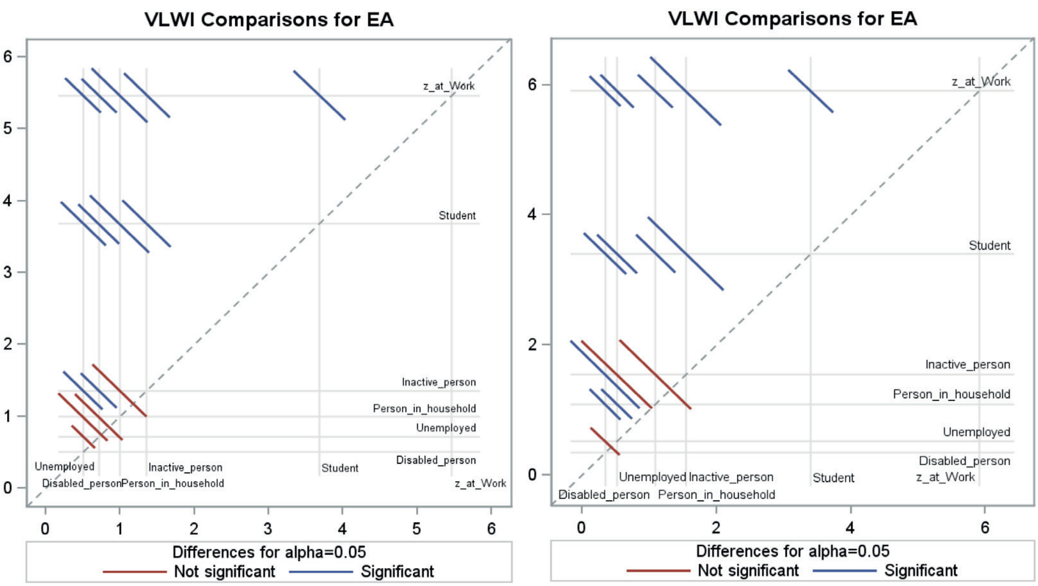
Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

not have a significant impact on the target variable at the significance level of 0.1. Only those variables listed in Table 1 were included in the logit model separately estimated for Slovakia and the Czech Republic. Table 2 confirms the statistical significance ($\alpha = 0.1$) of the impact of these variables on the target variable.

In both countries, economic activity is a key factor, that significantly affects VLWI. Other factor in the both countries, that fundamentally affects the probability that a person lived in a (quasi-)joblessness household in 2021 is the type of household. Through this variable, we confirm that not only the number of dependent children, but also the number of economically active members in the household has an impact on exclusion from the labor market. In the case of Slovakia, the third most important factor is a person's education, and in the Czech Republic it is age.

An analysis of least squares means performed using the LSMEANS statement in the SAS programming language showed, which categories of individual factors are the most risky and the least risky in terms of (quasi-)joblessness and between which pairs of categories of individual factors there are significant differences in the marginal means of the logit probability of living in a (quasi-)joblessness household. Based on the results of the analyses, we will assess the possibility of clustering the most similar categories into one newly created category and through this we will simplify following results.

Figure 1 Interval estimates of the LS means of the logit of (quasi-)joblessness depending on the status of economic activity (left Slovakia, right the Czech Republic)



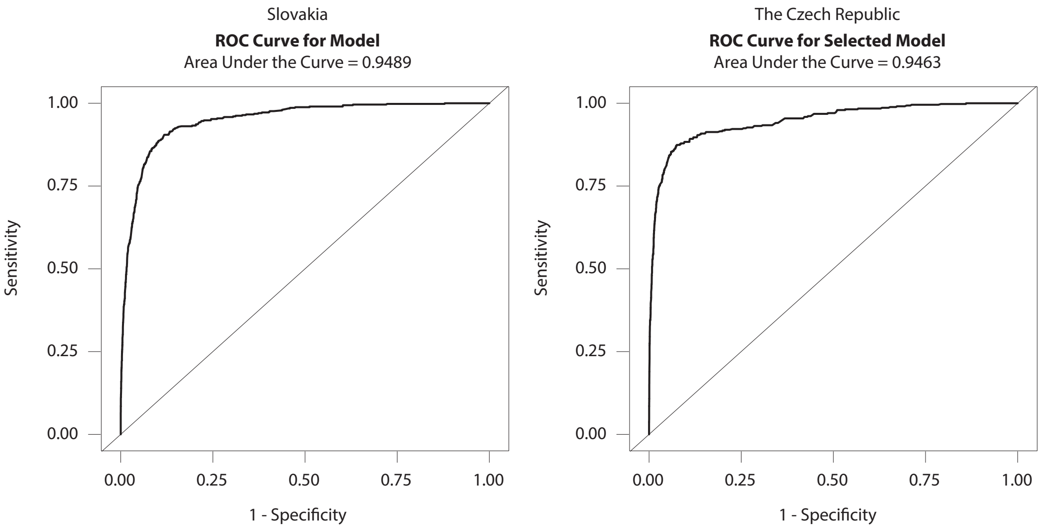
Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

In the case of several pairs of EA and Education factor categories, statistically significant differences (Figure 1, red lines) in the marginal means of the logit were not confirmed at the 0.05 significance level. In 2021, in both analyzed countries, there was no significant difference between disabled people and unemployed people and between inactive people and people in household (Figure 1). In the next analysis, we used clustered pairs of categories and denoted them DP-UP (Disabled Persons and Unemployed Persons) and IP-PH (Inactive Persons and Persons in Household). We proceeded similarly in the case of the Education factor, in which we clustered the ISCED 3 and ISCED 4-5 categories and the ISCED 6 and ISCED 7-8 categories, between which there was no significant difference. After this modification, the Education factor had 3 categories, namely ISCED 0-2, ISCED 3-5 and ISCED 6-8.

4.2 Quality of logit models and odds ratios estimation of (quasi-)joblessness in Slovakia and the Czech Republic

Adequacy and success of the model in prediction are confirmed by the association measures, namely Sommer's D, Goodman-Kruskal's gamma or the c statistic, whose values are relatively high. The comparison of concordant and discordant pairs also indicates good model quality.

Figure 2 ROC curve for the logit model of Slovakia and the Czech Republic



Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

The quality of the model is also confirmed by the area under the ROC curve (AUC) characterized by the c statistic, which is 0.9483 in Slovakia and 0.9465 in the Czech Republic (Figure 2).

Economic activity naturally has the greatest impact on the risk of (quasi-)joblessness. In both countries the most at risk are disabled and unemployed persons (DP-UP category). The odds of living in a (quasi-) joblessness household for persons living in Slovakia, who are disabled or unemployed are 127.661 times higher than for employed persons. This odd ratio is almost 2 times higher in the Czech Republic (*OR* = 242.54). Let us note that these results do not indicate that the probability of living in (quasi-) joblessness household in the Czech Republic is greater for disabled and unemployed persons than in Slovakia. This oddity compares the probability of living in a QJ household for disabled and unemployed persons in relation to the probability of living in a QJ household for employed persons. The stated result may be a consequence of the fact that employed persons in Slovakia have a higher probability to live (quasi-)joblessness household than in the Czech Republic. As shown in Figure 3a), in our case the reason for the higher odds ratio in the Czech Republic is precisely this fact.

Other high-risk economic activity statuses in both countries are otherwise inactive people and people in the household. Such people again have a higher odd in the Czech Republic (*OR* = 118.803) than in Slovakia (*OR* = 71.190). Johnston and McGauran (2018) or Šoltés et al. (2022) also confirmed that disabled and unemployed people are the most risky and employed people are the least risky in terms of exclusion from the labor market.

Table 3 Parameter estimates and odds ratios

Analysis of Maximum Likelihood Estimates and Odds Ratio Estimates							
Parameter		**Slovakia**			**Czech Republic**		
		Estimate	**p-value**	**OR**	**Estimate**	**p-value**	**OR**
Effect		–8.6759	<.0001	–	–7.6830	<.0001	–
EA	DP-UP	4.8494	<.0001	127.661	5.4913	<.0001	242.540
	IP-PH	4.2654	<.0001	71.190	4.7776	<.0001	118.803
	Student	1.8584	<.0001	6.413	2.5400	<.0001	12.678
	at_Work	.					
HT	1A_0Ch	2.9550	<.0001	19.203	2.1277	<.0001	8.395
	1A_1+Ch	2.4064	<.0001	11.093	2.6470	<.0001	14.111
	2A(1+R)	3.5837	<.0001	36.006	2.9903	<.0001	19.891
	2A_0Ch	1.3836	<.0001	3.989	0.9384	0.0012	2.556
	2A_1Ch	1.0929	0.0016	2.983	0.3359	0.2165	1.399
	2A_3+Ch	1.1684	0.0013	3.217	0.5148	0.1031	1.673
	Other_0Ch	1.1893	0.0005	3.285	0.7215	0.0164	2.058
	Other_1+Ch	1.0434	0.0011	2.839	–0.3654	0.3387	0.694
	2A_2Ch	.					
ED	ISCED 0-2	2.3241	<.0001	10.217	0.4981	0.0242	1.646
	ISCED 3-5	0.9585	0.0002	2.608	0.0305	0.8946	1.031
	ISCED 6-8	.					
MARITAL STATUS	Divorced	0.3873	0.1420	1.473	–0.3139	0.1987	0.731
	Never_married	0.4531	0.0091	1.573	0.5732	0.0044	1.774
	Widowed	0.1552	0.6854	1.168	–0.1051	0.8384	0.900
	Married	.					
AGE	30–40	0.5021	0.0141	1.652	0.3660	0.1205	1.442
	40–50	0.6038	0.0075	1.829	1.0203	<.0001	2.774
	50+	0.3309	0.1667	1.392	1.4256	<.0001	4.160
	–30	.					
URB	Intermediate	0.5285	0.0044	1.696	–0.3392	0.0414	0.712
	Sparse	–0.1988	0.2827	0.820	–0.4196	0.0129	0.657
	Dense	.					

Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

The second most affecting factor is the type of household. In both countries, in terms of household type, the most risky persons living in household type 2 adults with one person aged 65+. In Slovakia, the odd of living in a (quasi-)joblessness household for such a person was 36.006 times higher than for a person from a household of 2 adults with 2 dependent children. In the Czech Republic was this odd ratio at the level of 20. Since the work intensity of a household depends on the use of work potential by all adults in the household, in the case of households with one adult person of productive age (1A_1+Ch, 1A_0Ch, 2A(1+R)) there is a threat of very low work intensity per person , while in households with more adults, this risk is shared among more adults, and therefore it is natural that a person living in a household with more adults has a lower risk of living in a QJ household. This assumption is also confirmed by the results of our research shown in Table 3.

Based on education, in both countries, the lowest education level (ISCED 0-2) is the most risky and the tertiary level of education (ISCED 6-8) is the least risky. Odds ratios estimated for individual categories of other factors (Table 3) show that the lowest risk of living in a QJ household is in Slovakia and also in the Czech Republic for never married people, in Slovakia for people aged 40–50 and in the Czech Republic for people aged 50 +. In terms of urbanisation, these are intermediate areas in Slovakia and densely-populated area in the Czech Republic. Compared to odds ratios, the probabilities provide a better idea of the factor impact on the risk of living in a QJ household. Therefore, in the next parts of the paper, we will focus on the estimates of these probabilities.

4.2.1 Probabilities of (quasi-)joblessness in relation to individual factors

Figures 3a to 3f present the estimated probability of QJ in Slovakia and the Czech Republic for individual categories of factors. These probabilities were estimated based on the logit models presented in the previous part of the paper using the ESTIMATE statement in the SAS programming language.

Factor EA

Figure 3a confirms our previous finding that the most risky group of persons in both countries are disabled and unemployed persons, for which we estimated the probability of living in a QJ household at the level of 40.9% (33.7%–48.5%)[4] in Slovakia, or 39.5% (31.2%–48.5%) in the Czech Republic. The next most risky statuses of economic activity are inactive persons and persons in the household. Across these two statuses, we estimated the probability of QJ at the level of 27.9% (20.5%–36.7%) in Slovakia and at the level of 24.2% (17.9%–31.9%) in the Czech Republic. The lowest risk of exclusion from the labor market has employed persons for whom the probability of QJ does not exceed 1% in any of the considered countries. All the above results for individual economic activity statuses are estimated across all categories of individual factors included in the logit models. So, these are the mean QJ probabilities for persons with different economic activity statuses across all levels of education, types of households, marital statuses, age categories and urbanization. In this sense, it is necessary to interpret the QJ probabilities estimated in other parts of this subsection.

Factor HT

Figure 3c confirms the finding we arrived at based on the odds ratios (Table 3) that the greatest risk of QJ in 2021 was for persons living in households with two adults, at least one of whom is 65 or older. Figure 3c confirms (Table 3) that the lowest risk of QJ in 2021 was for persons who live in households of 2 adults of which one is aged 65 and older. In Slovakia, we estimated the probability for this category at the level of 36.9% (24.3%–51.6%) and the Czech Republic by approximately 6 pp lower. The second most risky category in Slovakia was a household of 1 adult, with a probability of 23.8%. In the Czech Republic were people from this type of household in the third place of risk (15.6%), because they were overtaken in the risk of QJ by persons from households of 1adult with at least 3 dependent children (23.7%). People from these 3 types of households, i.e. from households in which there is only 1 adult of productive age, had a significantly higher risk of QJ than persons from other types of households, in which the probability was below 5% in both countries.

Factor Education

The results of our analyses presented in Figure 3b revealed that in 2021 the Education factor had a different effect on the probability of QJ in Slovakia and the Czech Republic and showed that education

[4] In the full paper, after the point estimate of probability is present in parentheses the 95% confidence interval.

had a lower impact on the target variable in Slovakia than in the case of the Czech Republic. Although in both countries the most risky persons were those with the lowest level of education (in Slovakia 17.3%–28.4%; in the Czech Republic 6.2%–11.2%), in Slovakia the probability of living in QJ households for these persons was more than 10 pp higher than in the Czech Republic. In general, with the increasing education level of persons, the probability of the risk of QJ decreases. However, in the case of the Czech Republic non-significant difference ($p = 0.8946$) was confirmed between persons with higher secondary education or post-secondary education (ISCED 3-5: 5.4%) and tertiary education (ISCED 6-8: 5.3%). Economic activity, education and type of household impact risk of QJ to a greater extent than other factors, which is also confirmed by the fact that in the most risky categories of the following factors, the probability did not exceed 12% in any country.

Factor Age

Based on Figure 3d, age causes smaller disparities in the probability of QJ in Slovakia than in the Czech Republic. In Slovakia, the most risky persons are aged 40–50 (9.7%; 6.8%–13.8%), while in the Czech Republic they are aged over 50 (12.0%; 8.3%–17, 0%). While in the Czech Republic the probability of QJ in 2021 increased with increasing age, in Slovakia this pattern was disrupted by the age category over 50 (see Figure 3d).

Factor Marital_Status

The riskiest marital status in 2021 in terms of QJ was never married. In Slovakia, persons with this marital status had a probability of living in a QJ household of 9.4% (7.4%–11.8%) and in the Czech Republic 10.2% (8.0%–12.8%). Married persons (6.12%) were least at risk in Slovakia and divorced persons (4.5%) in the Czech Republic, which is quite strange, since Divorced status was the second most risky in Slovakia. In both countries, there was a statistically non-significant difference in the probability of QJ between widowed and married persons (Slovakia: $p = 0.6854$, Czech Republic: $p = 0.8384$).

Factor Urbanisation

In the Czech Republic, the level of urbanization clearly had the lowest impact on the target variable of all factors, so it was not clear in Slovakia (Figure 3f). In terms of the level of urbanization, in 2021 the least risky areas were sparse populated areas (in Slovakia 5.8% and Czech Republic 5.3%, respectively

Figure 3 Probabilities estimates of living in a QJ household depending on the status of economic activity (a), education (b), type of household (c), age (d), marital status (e) and urbanization (f) for Slovakia (left) and the Czech Republic (right) in 2021

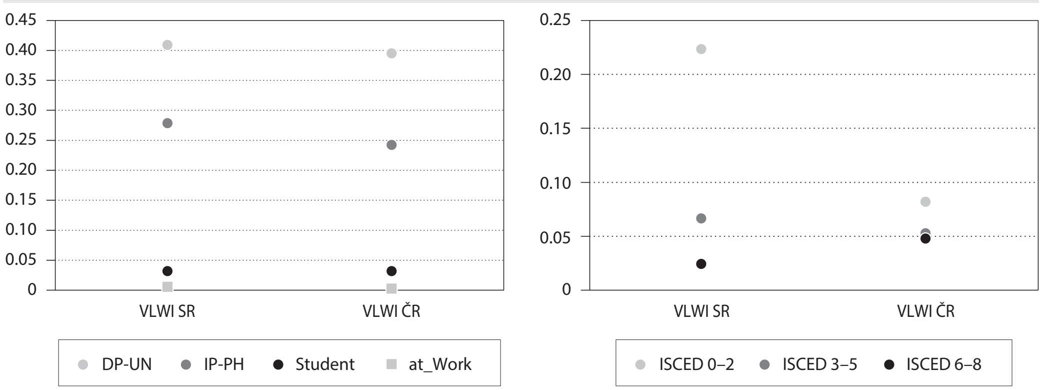
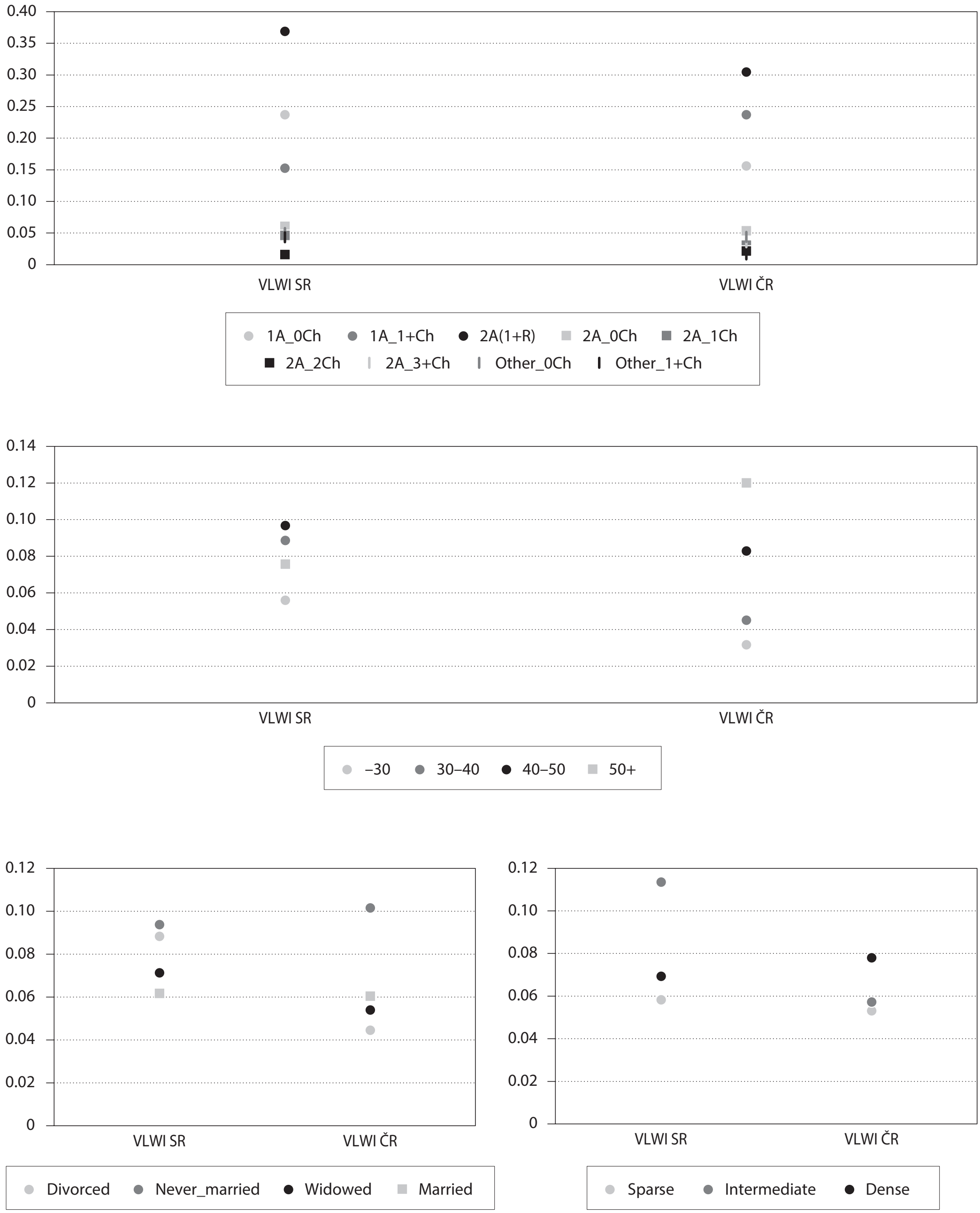


Figure 3 (continuation)



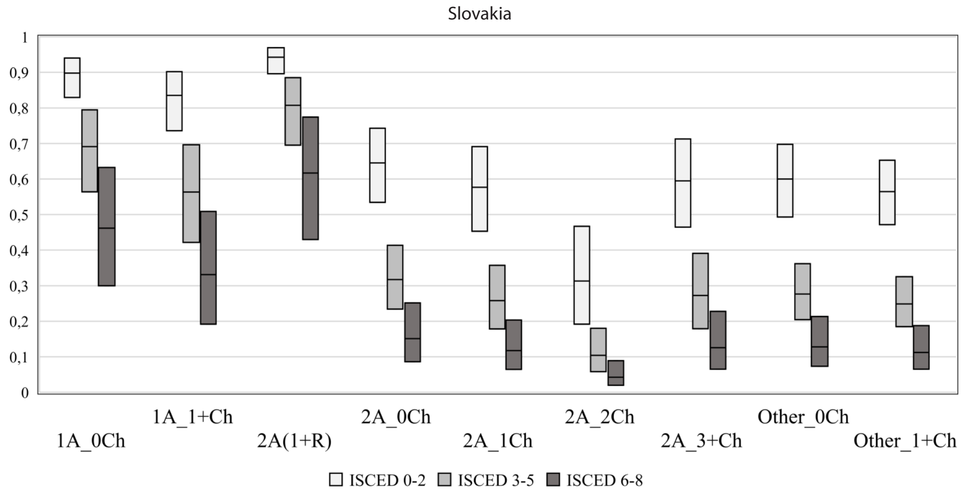
Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

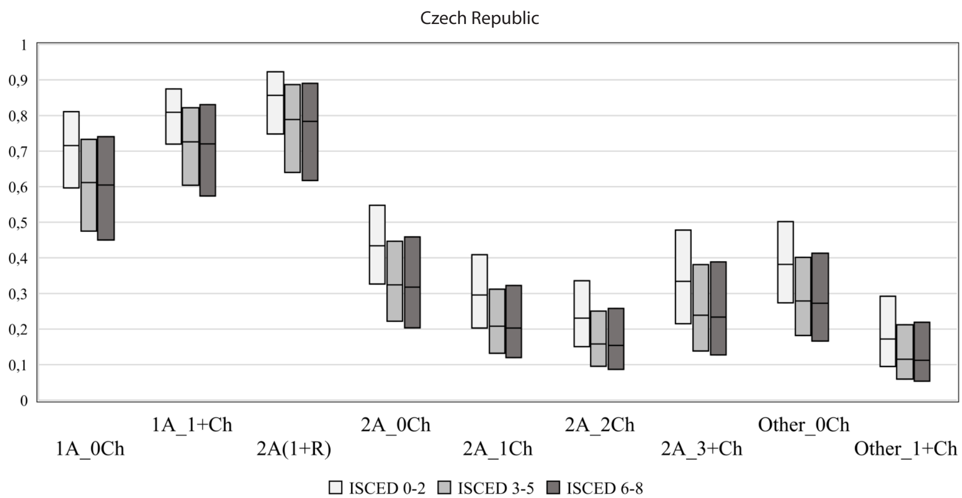
4.2%–8.0% and 3.7%–5.8%). While intermediate populated areas were the most risky in Slovakia (11.4%, 8.5%–15.0%), in the Czech Republic it was densely populated areas (7.9%; 5.6%–11.0%).

4.2.2 Probabilities of QJ for people profiles determined by their status of economic activity, education and type of household

In the following part of the paper, we estimate the probabilities of QJ depending on the interaction of the three most important factors, namely status of economic activity, type of household in which this person lives and his education. For a better overview, we present the results only for the two most frequent statuses of economic activity, namely the status "unemployed or disabled" and the status "employed", while at the same time it is the most risky and the least risky status. Figure 4 shows the probability estimates for unemployed or disabled people for different profiles of people determined by education and household type. Figure 5 shows estimates of such probabilities, but for people with the status "employed". Both

Figure 4 Point and interval probability estimates of QJ for unemployed and disabled people in 2021 depending on education and household type (Slovakia – left, Czech Republic – right)



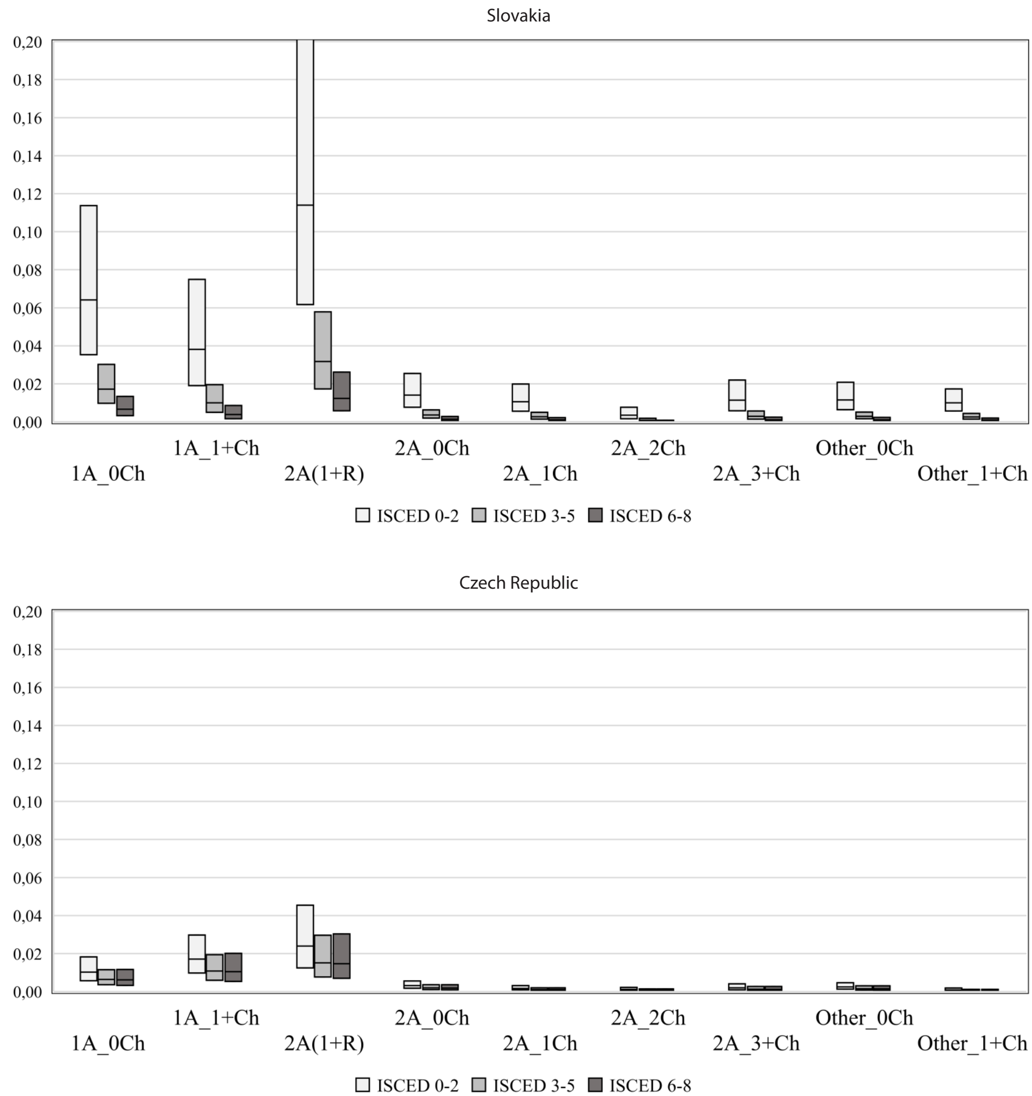


Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

figures (Figure 4 and Figure 5) provide a comparison of the relevant probabilities between Slovakia and the Czech Republic in 2021.

Among other things, the analyzes presented in the previous part of the article showed that the probability of living in a QJ household decreases with increasing education. Figure 4 confirms that this is the case in all types of households. In terms of a education we can observe more fundamental probability differences in Slovakia. This finding corresponds to the previous finding that in the Czech Republic, education has a smaller impact on the risk of living in a QJ household.

Figure 5 Point and interval probability estimates of QJ for employed persons in 2021 depending on education and type of household (Slovakia – left, Czech Republic – right)



Source: EU-SILC 2021 SO SR and CR, own processing in SAS EG

The highest probability of living in a household with very low work intensity among unemployed and disabled persons is observed in households with 2 adults where one is aged 65+. In 2021, this probability for persons with the lowest level of education (ISCED 0-2) was up to 94.3% (89.6%–96.9%), for persons with higher secondary or post-secondary education (ISCED 3-5) at the level of 80.8% (69.5%–88.5%) and for people with higher education at the level of 61.7% (43.0%–77.4%), if it is Slovakia, respectively 85.6% (74.8%–92.3%) for ISCED 0-2, 78.9% (64.0%–88.9%) for ISCED 3-5 and 78.4% (61.7 %–89.0%) for ISCED 6-8, if it is the Czech Republic.

Other results showed that people living in households of type 1A, 1A_1+Ch and 2A(1+R) had a significally higher probability of living in QJ households in 2021 than people from other types of households. This finding applies to both countries. There are no significant differences in the risk of QJ between people from other types of households. However, the exception is people living in a household of 2 adults with 2 dependent children, who have demonstrably the lowest risk of QJ in Slovakia. In the Czech Republic, people from the household type of 2 adults with 2 dependent children also had a relatively low risk of QJ, but not the lowest.

The probability of QJ for unemployed or disabled persons with low education (ISCED 0-2) was generally above 50% in Slovakia (the exception is the household type 2A_2Ch) and even above 75% in households 1A_1+Ch, 1A, 2A(1+R). In the Czech Republic, this probability for unemployed or disabled persons with the lowest level of education in households 1A_1+Ch, 1A, 2A(1+R) was over 60% in other types of households mostly in the range of 20%-50% (exceptions are household types 2A_2Ch and Other_1+Ch). Employed people (Figure 5) had a considerably lower probability. Even for persons with the lowest education in most types of households did not exceed 2%. The probability of living in a QJ household for employed persons exceeded 2% only in households types 1A_1+Ch and 2A(1+R) in the Czech Republic and in households types 1A_1+Ch, 1A and 2A(1+R) in Slovakia.

DISCUSSION AND CONCLUSION

The presented results are based on logit models and subsequent analysis of marginal means and contrast analysis of QJ in Slovakia and the Czech Republic. By (quasi-)joblessness (QJ) we understand the risk of living in a household that has a very low work intensity. In Slovakia and the Czech Republic, the probability of (quasi-)joblessness is primarily linked to economic activity and type of household, followed by education in Slovakia and age in the Czech Republic, while the authors Verbunt and Guio (2019) reached a similar conclusion for the countries of the European Union. Marital status and urbanization have a smaller, but still significant iimpact on (quasi-)joblessness.

Analyses of individual levels of factors revealed that unemployed and disabled people are most risky in both countries. In Slovakia, the odd of living in a (quasi-)joblessness household for people who are either unemployed or disabled is up to 127.661 times higher than for employed people, while in the Czech Republic this odds ratio is almost 2 times higher. In terms of household type, single-parent and multi-child households are the most risky, which was also confirmed by the findings of the authors Verbunt and Guio (2019) or the authors Filandri and Struffolino (2019), who dealt with another dimension of social exclusion, namely the risk of income poverty. In our analyses, we have revealed that the highest probability of living in a household with very low work intensity among unemployed and disabled people is observed in both countries in households with 2 adults, where one is aged 65+.

For unemployed and disabled people and individual types of households, the probability of very low work intensity decreases with an increase in educational level. However, the results of our analyses showed that education had a much greater impact on QJ in Slovakia in 2021 than in the Czech Republic. Even though we have confirmed in the case of both countries that the persons with the lowest level of education are most risky, in Slovakia the probability of living in QJ households for persons with the lowest educational level was approximately 14 pp higher than in the Czech Republic. The lowest level of education has proven to be a risk factor for poverty and social exclusion in other European Union countries as well, which is also confirmed by the authors Filandri and Struffolino (2019), Dudek and Szczesny (2021).

Our analyses show that the probability of living in a QJ household in the Czech Republic decreases with increasing age. In Slovakia, this pattern of dependence was disturbed by the age category 50+, for which we estimated a lower probability of QJ than in the age categories 30–40 and 40–50. In Slovakia, the most risky persons are aged 40-50 (9.7%), in the Czech Republic it is people over 50 (12.0%). In terms of marital status, never married persons were the most risky in both countries. In Slovakia, never married persons had a probability of living in QJ at the level of 9.4% and in the Czech Republic we estimated this probability at the level of 10.2%. Unexpected was the finding that in Slovakia the people who lived in marriage were the least risky, while in the Czech Republic it was divorced people. In terms of urbanisation, intermediate populated areas were the most risky in Slovakia (11.4%) and densely populated areas in the Czech Republic (7.9%). The stated findings are relatively unexpected, because in terms of other dimensions of poverty and social exclusion (income poverty or material deprivation), sparsely populated areas are the most risky, which was confirmed by e.g. Weziak-Bialowolska (2016).

The above conclusions were obtained by contrasting analysis and subsequent estimation of the probability depending on individual factors, while the impact of other factors was fixed. In addition, we estimated the probabilities of QJ in the interaction of the three most important factors, namely economic activity, type of household and education, while the other relevant factors (age, marital status and degree of urbanization) were fixed. For a better overview, we focused only on the two most frequent statuses of economic activity, namely the status "unemployed or disabled" and the status "employed", while this is the most risky and the least risky status of economic activity in terms of (quasi-)joblessness. We estimated the probabilities of QJ for 54 groups of people in Slovakia, depending on the 2 statuses of economic activity, 9 household types and 3 categories of education ($2 \times 9 \times 3 = 54$) and we also estimated the probabilities of QJ for 54 groups of people in the Czech Republic. Our analyses showed that unemployed and disabled people clearly have a higher risk of living in a QJ household in all education levels and in all types of households. People from households with only 1 adult of productive age had a higher risk of QJ than people from other types of households, in all education levels, with more visible differences among unemployed or disabled people than in the case of employed people. For people who are from these most risky types of households, who are simultaneously unemployed or disabled and have the lowest level of education (ISCED 0-2), we estimated the probability of QJ above 75% in Slovakia and above 60% in the Czech Republic.

In conclusion, it should be said that the paper provides a limited view of the risk of living in a (quasi-) joblessness household due to the fact that we analyzed the impact of only selected factors. In addition, we focused on only one of the dimensions of poverty and social exclusion, which opens up space for further research and the expansion of our analyses by other aspects of social exclusion, such as the risk of poverty or material deprivation. Despite some limitations, the paper reveals the most risky groups of people in terms of very low work intensity in the Slovakia and Czech Republic. Social policy should focus primarily on the mentioned groups of people.

ACKNOWLEDGMENTS

This work was supported by the VEGA projects *The Impact of Inflation on Poverty and Social Exclusion in Slovakia and the EU* (No. 1/0285/24) and *The impact of the COVID-19 crisis on business demography and employment in the Slovak Republic and the EU* (No. 1/0561/21).

References

ALLISON, P. D. (2012). *Logistic Regression using SAS. Theory and Application.* 2nd Ed. Cary, NC: SAS Institute Inc.

ATKINSON, A. B., GUIO, A. C., MARLIER, E. (2017). (eds.) *Monitoring social inclusion in Europe.* Luxembourg: Publications Office of the European Union.

BATZER, E. (2020). *Contrasts and Linear Hypothesis Tests* [online]. <https://www.evanbatzer.com/posts/contrasts>.

BRADSHAW, J., WILLIAMS, J., LEVITAS, R., PANTAZIS, C., PATSIOS, D., TOWNSEND, P., MIDDLETON, S. (2000). The relationship between poverty and social exclusion in Britain. In: *26th General Conference of The International Association for Research in Income and Wealth*, Cracow, Poland.

BROWN, U., LONG, G. (2018). Poverty and welfare [online]. In: *Social Welfare*, Routledge, 19–34. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315202686-3/poverty-welfare-usha-brown-gil-long>.

DE GRAAF-ZIJL, M., NOLAN, B. (2011). Household joblessness and its impact on poverty and deprivation in Europe [online]. *Journal of European Social Policy*, 21(5): 413–431. <https://doi.org/10.1177/0958928711418854>.

DUDEK, H., LISICKA, I. (2013). Determinants of poverty – binary logit model with interaction terms approach. Econometrics [online]. *Ekonometria. Advances in Applied Data Analytics*, 3(41): 65–77. <https://www.researchgate.net/publication/270285045>.

DUDEK, H., SZCZESNY, W. (2021). Multidimensional material deprivation in Poland: a focus on changes in 2015–2017 [online]. *Quality & Quantity*, 741–763. <https://doi.org/10.1007/s11135-020-01024-3>.

DUIELLA, M., TURRINI, A. (2014). Poverty developments in the EU after the crisis: a look at main drivers [online]. *ECFIN Economic Brief*, 1–10. <https://data.europa.eu/doi/10.2765/72447>.

EUROSTAT. (2021). *Glossary: Persons living in households with low work intensity* [online]. <https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Glossary:Persons_living_in_ households_with_low_work_intensity>.

EUROSTAT. (2022a). *Risk-of-poverty in very low work intensity households at 64%* [online]. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220804-1>.

EUROSTAT. (2022b). *Living conditions in Europe – poverty and social exclusion* [online]. <https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Living_conditions_in_Europe-_poverty_and_social_exclusion&oldid=575050>.

FILANDRI, M., STRUFFOLINO, E. (2019). Individual and household in-work poverty in Europe: understanding the role of labor market characteristics [online]. *European Societies*, 21(1): 130–157. <https://doi.org/10.1080/14616696.2018.1536800>.

GALLARDO, M. (2020). Measuring vulnerability to multidimensional poverty [online]. *Social Indicators Research*, 148(1): 67–103. <https://doi.org/10.1007/s11205-019-02192-y>.

GAO, M., LI, L., GAO, Y. (2022). *Statistics and Analysis of Targeted Poverty Alleviation Information Integrated with Big Data Mining Algorithm* [online]. Security and Communication Networks. <https://doi.org/10.1155/2022/1496170>.

GUIO, A. C., MARLIER, E., VANDENBROUCKE, F., VERBUNT, P. (2022). *Differences in Child Deprivation Across Europe: The Role of In-Cash and In-Kind Transfers* [online]. Child Indicators Research, 1–26. <https://doi.org/10.1007/s12187-022-09948-y>.

JOHNSTON, H., MCGAURAN, A. M. (2018). *Low Work Intensity Households and the Quality of Supportive Services: Detailed Research Report.* Dublin: National Economic and Social Council.

LAFUENTE, J. Á., MARCO, A., MONFORT, M., ORDÓÑEZ, J. (2020). Social exclusion and convergence in the EU: An assessment of the Europe 2020 strategy [online]. *Sustainability*, 12(5): 1843. <https://doi.org/10.3390/su12051843>.

LITTELL, R. C., STROUP, W. W., FREUND, R. J. (2010). *SAS for Linear Models.* 4th Ed. Cary, NC: SAS Institute Inc.

NOLAN, B., MARX, I. (2009). *Economic inequality, poverty, and social exclusion* [online]. The Oxford handbook of economic inequality, 315–341. <https://doi.org/10.1093/oxfordhb/ 9780199606061.013.0013>.

OECD. (2009). *Is work the best antidote to poverty?* [online]. Chapter 3, Employment Outlook, Paris: Organisation for Economic Co-operation and Development. <https://www.oecd.org/els/emp/ 45219514.pdf>.

RAVALLION, M. (2020). On measuring global poverty [online]. *Annual Review of Economics*, 12: 167–188. <https://doi.org/10.1146/annurev-economics-081919-022924>.

RUSNAK, Z. (2012). Logistic regression model in poverty analyses [online]. *Econometrics*, 35: 9–23. <https://www.ceeol.com/search/article-detail?id=119600>.

SACCONE, D., DEAGLIO, M. (2020). Poverty, emergence, boom and affluence: a new classification of economies [online]. *Economia Politica*, 37(1): 267–306. <https://doi.org/10.1007/s40888-019-00166-4>.

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

SCHAD, D. J., VASISHTH, S., HOHENSTEIN, S., KLIEGL, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial [online]. *Journal of Memory and Language*, 110: 104038. <https://doi.org/10.1016/j.jml.2019.104038>.

STANLEY, J., HENSHER, D. A., STANLEY, J., CURRIE, G., GREENE, W. H., VELLA-BRODRICK, D. (2011). Social exclusion and the value of mobility [online]. *Journal of Transport Economics and Policy (JTEP)*, 45(2): 197–222. <http://www.jstor.org/stable/23072175>.

SUZUKI, M., TANIGUCHI, T., FURIHATA, R., YOSHITA, K., ARAI, Y., YOSHIIKE, N., UCHIYAMA, M. (2019). Seasonal changes in sleep duration and sleep problems: A prospective study in Japanese community residents [online]. *PLoS One*, 14(4): e0215345. <https://doi.org/10.1371/journal.pone.0215345>.

ŠOLTÉS, E., HURBÁNKOVÁ, Ľ., KOTLEBOVÁ, E., ŠOLTÉSOVÁ, T., VOJTKOVÁ, M. (2018). *Chudoba a sociálne vylúčenie v EÚ a v SR: v kontexte stratégie Európa 2020.* Pardubice: Univerzita Pardubice, Fakulta ekonomicko-správní. ISBN 978-80-7560-186-5

ŠOLTÉS, E., KOMARA, S., ŠOLTÉSOVÁ, T. (2022). Exploration of poverty and social exclusion of Slovak population via contrast analysis associated with logit models [online]. *Quality & Quantity*, 1–27. <https://doi.org/10.1007/s11135-022-01573-9>.

ŠOLTÉS, E., VOJTKOVÁ, M. (2018). Work intensity in Slovakia in relationship with socio-economic characteristics of households [online]. *Entrepreneurial Business and Economics Review*, 6(1): 115–134. <https://doi.org/10.1007/s11135-022-01573-9>.

VERBUNT, P., 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 [online]. *Social Indicators Research*, 144(2): 827–868. <https://doi.org/10.1007/s11205-018-2021-1>.

VLAČUHA, R., KOVÁČOVÁ, Y. (2021). *EU SILC 2020 Indikátory chudoby a sociálneho vylúčenia.* Bratislava: ŠÚ SR.

WANG, B., WU, P., KWAN, B., TU, M. X., FENG, C. (2018). Simpson's paradox: examples [online]. *Shanghai Archives of Psychiatry*, 30(2): 139. <http://dx.doi.org/10.11919/j.issn.1002-0829.218026>.

WEZIAK-BIALOWOLSKA, D. (2016). Spatial variation in EU poverty with respect to health, education and living standards [online]. *Social Indicators Research*, 125(2): 451–479. <https://doi.org/10.1007/s11205-014-0848-7>.