Overeducation of tertiary education graduates in Slovakia: relationship with the field of study

Dávid Martinák
Institute of Economic Research
of Slovak Academy of Sciences

The main objective of the analysis is to explore the role of the field of study in determining the overeducation of tertiary education graduates. The individual-level analysis of overeducation determinants is performed on Slovak Labour Force Survey. We use a probit model with sample selection to take into account the selection of graduates into employment. The results suggest a strong relationship between study fields and overeducation risk. In particular, graduates in ICT; Arts and Humanities; and Education are less likely to become overeducated. The opposite is true for graduates in Services. Furthermore, we found evidence that some field of study effects are gender-specific. We also explore the role of labour shortages as an important channel through which field of study affects overeducation risk. Finally, based on observed empirical results implications for educational and labour market policies are suggested.

L'obiettivo principale dell'analisi è esplorare il ruolo dell'ambito di studio nel determinare il fenomeno dell'overeducation dei laureati. L'analisi a livello individuale dei fattori determinanti la sovra istruzione viene eseguita sulla base della Slovak Labour Force Survey. Viene adottato un modello probit correggendo per la possibile selezione dei laureati nel mercato del lavoro. I risultati suggeriscono una forte relazione tra ambiti di studio e il rischio di sovra istruzione. In particolare, i laureati in ICT, arte e scienze umane e istruzione sono meno a rischio di overeducation. Per i laureati nel settore dei servizi è vero il contrario. Inoltre, per alcuni ambiti di studio troviamo evidenza di effetti gender-specific. Viene anche indagato il ruolo della carenza di manodopera nell'influenzare il rapporto tra ambito di studio e rischio di sovra istruzione. Infine, sulla base dei risultati empirici osservati, si forniscono indicazioni per le politiche educative e del mercato del lavoro.

DOI: 10.1485/2532-8549-202003-7

Citazione
Martinák D. (2020), Overeducation of tertiary education graduates in Slovakia: relationship with the field of study, *Sinappsi*, X, n.3, pp.93-110

Key words
Tertiary education
Overeducation
Field study

Parole chiave
Educazione terziaria
Overeducation
Ambito di studio

1. Introduction

Overeducation refers to a situation when an individual works in a job with lower educational requirements than his obtained highest level of education. Terms like overschooling, overqualification, or more general qualification mismatch, and vertical qualification mismatch are also often used to refer to this phenomenon. The issue of overeducation has attracted the attention of researchers for quite a long time, while many research papers have detected negative effects of overeducation for both individuals and society. On the individual level, overeducation might harm earnings (Duncan and Hoffman 1981; Sicherman 1991) and job satisfaction (Hersch 1991) of workers compared

The work was supported by VEGA Grant Agency in the frame of the project VEGA no. 2/0182/17.

to their well-matched counterparts. In macro-level terms, national welfare may be potentially lowered by the underutilization of skills of overeducated workers (McGuinness 2006) since skill mismatches harm labour productivity via the allocative efficiency channel (McGowan and Andrews 2015b).

With the development of overeducation literature, more attention has been paid to the role of the field of study as a factor affecting the risk of overeducation. According to economic literature, there are several possible channels through which field of study may affect the labour market prospects of an individual. Fields of study can be understood as different stocks of human capital, differentially valued by employers (Ortiz and Kucel 2008). From a credentialism theory perspective, scientific and technical disciplines may be more selective and thus signalling a higher level of ability and motivation (Barone and Ortiz 2011). According to Reimer *et al.* (2008), educational expansion may be associated with a decline in mean ability in less academically challenging fields like humanities and social sciences. As a result of tertiary education expansion, a larger proportion of less academically successful students pursue their studies in tertiary education and these students often end up in less academically challenging fields. Consequently, the signal value of corresponding degrees may be lowered resulting in worse labour market prospects of graduates in these fields. Another aspect of the field of study impact may lie in the fact that some fields may be more occupationally oriented (Health and Welfare) than others (Social sciences, Humanities, Services). Therefore, graduates in more occupationally oriented fields do not face such a competition of graduates from other fields (Ortiz and Kucel 2008). Moreover, graduates in narrowly defined study fields may have better information regarding their job alternatives than their counterparts from soft study fields (Kucel and Byrne 2008) and thus may find an appropriate job more easily. This information asymmetry arises since a possible set of jobs related to hard fields of study is better defined and therefore the personal networks are more productive. Barone and Ortiz (2011) point out that it would be naive to expect that a rapid expansion of diplomas mechanically translates into overeducation. Much depends on the corresponding labour market opportunities for graduates. According to Croce and Ghignoni (2012) there is an inverse relationship between relative wages of university graduates and overeducation incidence in Europe suggesting an impact of demand for skilled labour. Montt (2015) used the concept of field saturation to capture demand-side factors at the level of each field of study. The reason behind is that fields of study may differ in terms of availability of jobs and job opportunities. For graduates in fields that experience higher levels of saturation, the priority might be to find work in the field, even if that means accepting a job with lower qualifications. Recent research in this area also shows that the effect of a field of study on the likelihood of becoming overeducated does not have to be homogenous, it may vary among genders. Rossen *et al.* (2019) identified several possible channels through which gender-specific field of study effects may affect the probability of overeducation. Men and women within the same field might have different preferences in terms of job attributes. Another possible channel is gender-specific discrimination. In other words, it may be harder for women or men to get a job as a consequence of gender stereotypes in some jobs.

The purpose of the present study is to explore the role of the field of study in determining overeducation of tertiary education graduates in Slovakia. Although a wide spectrum of literature has addressed the questions regarding the causes of overeducation only little is known about the determinants in Slovakia. Someone may argue that Slovakia is not a learning case study for an international audience, however, there are several reasons why the present study may be of interest even in international terms. Firstly, according to available sources, overeducation of tertiary-educated workers in Slovakia is high compared to other EU countries (Machlica *et al.* 2017) and is increasing over time (Eurostat 2019). According to Eurostat's experimental statistics on skills mismatches, Slovakia has experienced the highest increase in overeducation among EU countries between the years 2008 and 2019. Secondly, similarly to other transition economies in the region, Slovakia has experienced a significant increase in the supply of high-skilled workers since the early 90s associated with a significant expansion of tertiary education. The size of the active tertiary educated population has increased by more than 160% since 2000 (Slovak Statistical Office 2019). An increasing number of tertiary education graduates has raised several questions regarding their labour market allocation. Some studies suggest misallocation of the high-skilled labour

force (McGowan and Andrews 2015a) and the presence of information asymmetry among young students (Machlica *et al.* 2017) in Slovakia.

On the other hand, the labour market situation in Slovakia expressed by all the usually considered indicators has strongly improved over the last 10 years. According to Eurostat statistics, the unemployment rate dropped to historically low levels in 2019 (5.8%) and continued to decrease further until the onset of the Covid pandemic. Similarly, the unemployment rate of tertiary education graduates decreased to 2.5% in 2019. Also, the employment rate of tertiary educated active population has reached its highest level since 2009 (80.6%). Moreover, strong employment growth and the booming economy have brought labour shortages in some areas (OECD 2019). Although the demand for highly skilled workers has grown in the last decades, the increase appears to be lower compared to both OECD and countries in the region. The share of high-skilled jobs on total employment has increased by 4.5 pp. between the end of 1990s and 2015 compared to more than 7 pp. in Czech Republic and 7.5 pp. in Hungary (OECD 2017).

Many indicators show a low quality of tertiary education in Slovakia, whereas several studies emphasize the relationship between the quality of educational institutions and the probability of overeducation (McGuinness 2003; Ordine and Rose 2009). Slovak tertiary education is internationally rated as one of the weakest among EU countries and also among countries in the region (U21 2020). The number of top-quality publications per researcher is one of the lowest in the OECD (Machlica *et al.* 2017). Moreover, among OECD countries, Slovakia has the second-highest share of tertiary students abroad, with the most successful upper-secondary graduates studying abroad (Martinák and Varsík 2020). Another issue of Slovak tertiary education system is its weak link with the world of work, with most institutions organised as traditional academically oriented universities (Machlica *et al.* 2017). Similarly to many European countries, tertiary education is heavily subsidized in Slovakia and therefore from the policy perspective it is important to know whether educational investments are allocated effectively. The above-mentioned reasons make Slovakia a natural object for further investigation in this area.

To explore the role of the field of study in determining overeducation risk we employ micro-level analysis using Slovak Labour Force Survey (LFS). The present study is focused on young tertiary education graduates. In order to provide robust findings several definitions of overeducation are considered in the main analysis. In summary, the main contributions of the paper are the following: (1) improvement of the knowledge about the role of the field of study in overeducation incidence in Slovakia, (2) providing robust results by including sensitivity analysis, (3) addressing sample selection bias by employing probit model with sample selection, and (4) taking labour market situation into account.

2. Data and methods

Data

Slovak Labour Force Survey (LFS) is used as a main data source in the analysis. Compared to other considered sources it provides sufficient sample size even at the level of tertiary education graduates that are under our main focus. The main analysis is performed on the 2019 wave of LFS. Similarly to Rossen *et al.* (2019), we restrict our sample to tertiary-educated individuals. Armed forces occupations are excluded from the analysis. Furthermore, we include only individuals holding Bachelor and Master degree (ISCED 6 - 7) that are employed or non-employed and are less than 40 years old. Focusing on younger individuals is also in line with the research intention since we expect the relationship between education and labour market performance is very important especially at earlier career stages. We exclude graduates holding a PhD degree due to their small share in the labour force. Including them would often result in their unjustifiably high rates of overeducation when considering realized matches approaches. We exclude also individuals who declare being "pupil, student, further training, unpaid work experience" or "in retirement or early retirement or has given up business" as their main labour status. The restricted sample contains 4,688 observations with weighted proportions of 81% employed, 3% unemployed, and 15% inactive individuals. Due to the rotating sampling design, an individual may appear up to four times (quarters) in the yearly sample. Since an individual may occur up to four times in the 2019 sample, observations cannot be interpreted as inde-

pendent. Therefore, the individual-clustered standard errors are used in the regression analysis[1].

In order to capture the effect of labour shortages, we utilize the job vacancy data from the Slovak Central Office of Labour, Social Affairs and Family (UPSVAR) that contain the information about published job vacancies. The data provide useful information on the number of job vacancies, place of work, the ISCO classification of the job vacancy, and qualification requirements. Data contain information about approx. 261,000 job vacancies reported in 2019.

Methodological issues

Caroleo and Pastore (2013) pointed out three possible sources of bias related to the analysis of overeducation: endogeneity, measurement error and sample selection. Although their analysis is focused on the impact of overeducation on earnings, their reasoning is considerably relevant also when analysing determinants of overeducation what the case of this paper is.

First, endogeneity issues are related to unobserved characteristics of individuals that may differ in the level of endowed ability, skills or motivation. In order to avoid omitted variable bias, several studies tried to include skills measurements when analysing overeducation. As far as we know there is no data containing measurement of an individual´s ability or skills that involves a sufficiently large sample of tertiary education graduates in Slovakia.

Second, measurement issues arise since there is not any widely recognized measurement of overeducation in the literature. Several studies suggest that overeducation analysis is highly sensitive to the choice of measurement method (e.g. Rossen *et al.* 2019). To provide robust analytical findings, we employ a sensitivity analysis considering various overeducation indicators. All the indicators compare individuals' highest attained education with educational requirements of a particular job or occupation (See Model and variable description). However, there are several possible ways to measure occupational or job requirements. Leuven and Oosterbeek (2011) summarized job requirement measures into three broader categories: self-assessment, job analysis and realized matches benchmarks. Measures based on self-assessment rely on questions that directly ask workers on the schooling requirements of their job. Job analysis approach uses the required level of education based on information contained in occupational classifications. The realized matches approach derives educational requirements from the distribution of education (or schooling) among all workers in the same occupational category. LFS data allow to apply two measures of job requirements: job analysis and realized matches benchmarks.

Third, it may be reasonable to assume that the selection into employment is not random (Nicaise 2001). Moreover, since educational mismatch appears first of all in the form of a higher probability of non-employment and only then in a higher probability of overeducation (Caroleo and Pastore 2013), it can't be assumed that unobservables determining selection into employment are independent of those determining the probability of overeducation. Individuals with the same level of education, experience and other observed characteristics may differ in unobserved characteristics like personal motivation that presumably affect both the selection into employment and the probability of overeducation. Under these circumstances, estimation performed on a sample of employed individuals may suffer from sample selection bias. Ignoring selection bias will result in biased estimates and possibly in underestimation of the overeducation risk in study fields with higher non-employment. To address sample selection issues, we use the probit model with sample selection where we take into account the selection of graduates into employment. Our strategy is similar to Rossen *et al.* (2019) or Ghignoni and Verashchagina (2014) who use a two-step Heckman approach to control for self-selection into employment.

Model and variable description

Since it is assumed that there is a relationship between the unobserved error that affects whether an individual is overeducated and the unobserved error that affects whether an individual is employed or not, estimates gained using the simple probit model will be inconsistent. In other words, the unobserved factors that affect overeducation probability may be correlated with

1 In order to provide a sensitivity check, the estimates were also performed on first-quarter sample in which individuals occur only once. The obtained results are very similar from what is reported in the result section.

Table 1
Overview of overeducation indicators used in the analysis

Indicator	Description
Overeducation 1	An individual is considered as overeducated if his obtained years of education are higher than average years of education of all workers within the occupational group (2-digit ISCO) plus 1 standard deviation.
Overeducation 2	Similarly to Rossen *et al.* (2019) individual is considered as overeducated if his obtained years of education are higher than the 80th percentile of years of education within the occupational group (2-digit ISCO).
Overeducation 3	An individual is considered as overeducated if his obtained highest level of education exceeds the modal level of education within each occupational group (2-digit ISCO).
Overeducation 4	An individual is considered as overeducated if he holds tertiary education degree and works in occupations defined at 1-digit ISCO level as 4 and higher.

factors that affect the employment status of an individual. To take into account the selection of individuals into employment, we utilize a probit model with sample selection introduced by Van Praag and Van de Ven (1981). This method is analogous to Heckman's correction but for probit analysis since the dependent variable in both the outcome and selection equation is binary.

The dependent variable in the outcome equation is a binary overeducation indicator. In order to address the early-mentioned measurement issues and thus provide robust empirical results, we employ four strategies to measure overeducation previously used in literature (Table 1). Since LFS does not contain questions on self-assessment of occupation´s educational requirements, only realized matches (Overeducation 1 - 3) and job analysis approaches (Overeducation 4) may be considered to measure overeducation. Despite the different measurement methods, there is a positive correlation between all the considered indicators[2]. Educational benchmarks related to the indicators 1-3 were evaluated for the whole considered period (2014-2019). Since the "overeducation vs. upgrading" debate is not the subject of the study, the educational requirements of each occupation were not allowed to vary over time.

Since our main research interest focuses on the relationship between the field of study and overeducation, field of study dummies are included in the model. LFS provides information on the study field based on ISCED Fields of Education and Training classification (ISCED-F 2013) that includes 11 main study field categories out of which 10 are relevant at the tertiary education level in Slovakia. Business, Administration and Law was chosen as a reference category in the model due to his close to the median position in the distribution of mean overeducation incidence in all the considered indicators. As we indicated in the previous chapter, there are numerous channels through which field of study may affect the risk of overeducation. However, to avoid multicollinearity problems, it is not possible to incorporate field-specific indicators in the model specification.

To reveal the role of labour supply and demand in determining the overeducation risk the labour shortage concept is introduced. As a proxy for the labour shortage, we calculated the job vacancy rate that varies across regions and study fields. The regional variation in job vacancy rate allows for the inclusion in the model estimation together with the field of study dummies and thus separate the effect of labour shortages on overeducation. For more details see Appendix 2.

In addition to the field of study dummies and labour shortage index in outcome equation, we also control for several personal characteristics (gender, age, highest attained educational level, and region of residence), job-related characteristics (job tenure, full-time/part-time distinction), and household characteristics (presence of at least one unemployed person in the household). To explore possible gender-specific effects, we include also the interaction terms of gender with the field of study dummies and with the presence of unemployed persons in the household. Since due to collinearity reasons labour shortage indicator and regional dummies are mutually exclusive for estima-

2 The highest correlation (0.89) occurs between Overeducation 1 and Overeducation 3, and contrariwise, the lowest correlation (0.47) is between Overeducation 2 and Overeducation 3.

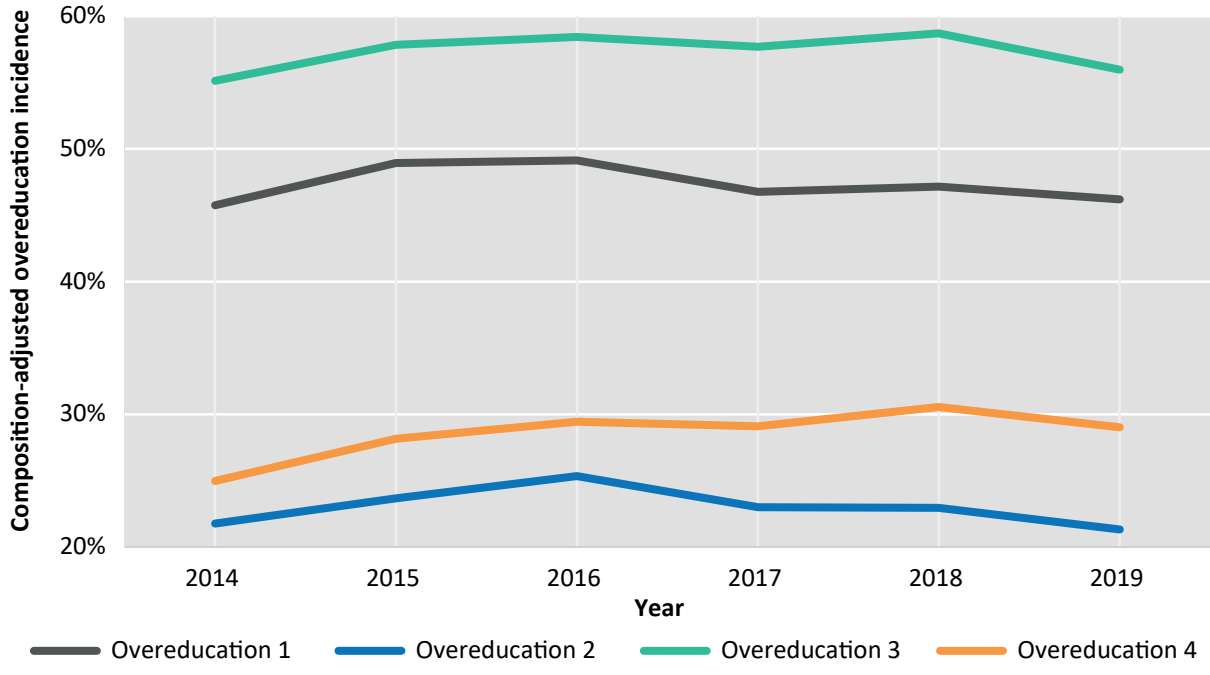
tion purposes, we estimated two model specification for each overeducation indicator. Consequently, eight outcome equation specifications are estimated.

The dependent variable in the selection equation is a binary variable where 1 means that an individual is employed and 0 not employed (unemployed or inactive). Selection equation contains the same independent variables as the outcome equation with exception of job-related characteristics. We include also several additional household characteristics that presumably have an impact on the probability of employment but not on the probability of overeducation. The household characteristics include the presence of at least one child aged less than 3 years in the household and the presence of a partner in the same household. Moreover, interactions of these household characteristics with gender are also included. The inclusion of additional variables in the selection equation is important for estimation purposes. It is often suggested for the model to be identified that the selection equation should contain at least one variable not included in the outcome equation. However, one may doubt the exclusion restriction itself and also the reasons for excluding some household characteristics and not the remaining (presence of an unemployed or inactive person in the household) from the outcome equation. Therefore, the validity of the exclusion restriction was checked following the strategy of Ghignoni and Verashchagina (2014) (see Appendix 2). The final estimation results are provided in Appendix 1.

3. Results

Figure 1 shows that the overall overeducation incidence varies significantly based on the chosen educational benchmark. The benchmarks settled by mean (Overeducation 1) and modal (Overeducation 3) approaches are much stricter than those settled by 80th percentile (Overeducation 2) and ISCO-ISCED approach (Overeducation 4). These differences support the consideration of the need for sensitivity analysis to provide robust results. Furthermore, also the evolution of overeducation over time slightly differs according to the chosen overeducation indi-

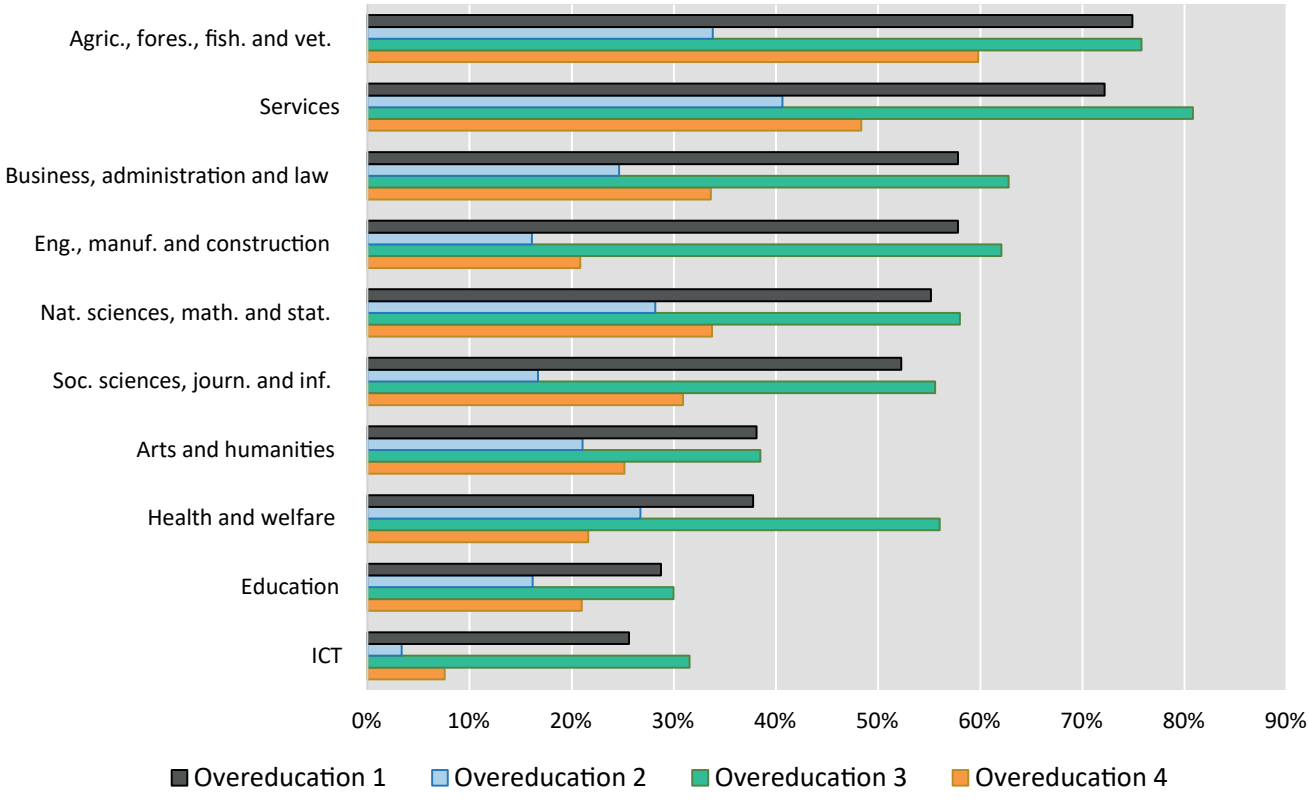
Figure 1
Evolution of overeducation in Slovakia



Notes: the composition adjustment of the average overeducation incidence holds the shares of the separate demographic groups defined by age, sex, and education constant over time.

Source: calculations based on LFS

Figure 2
Incidence of overeducation among fields of study in Slovakia



Source: calculations based on LFS (2019)

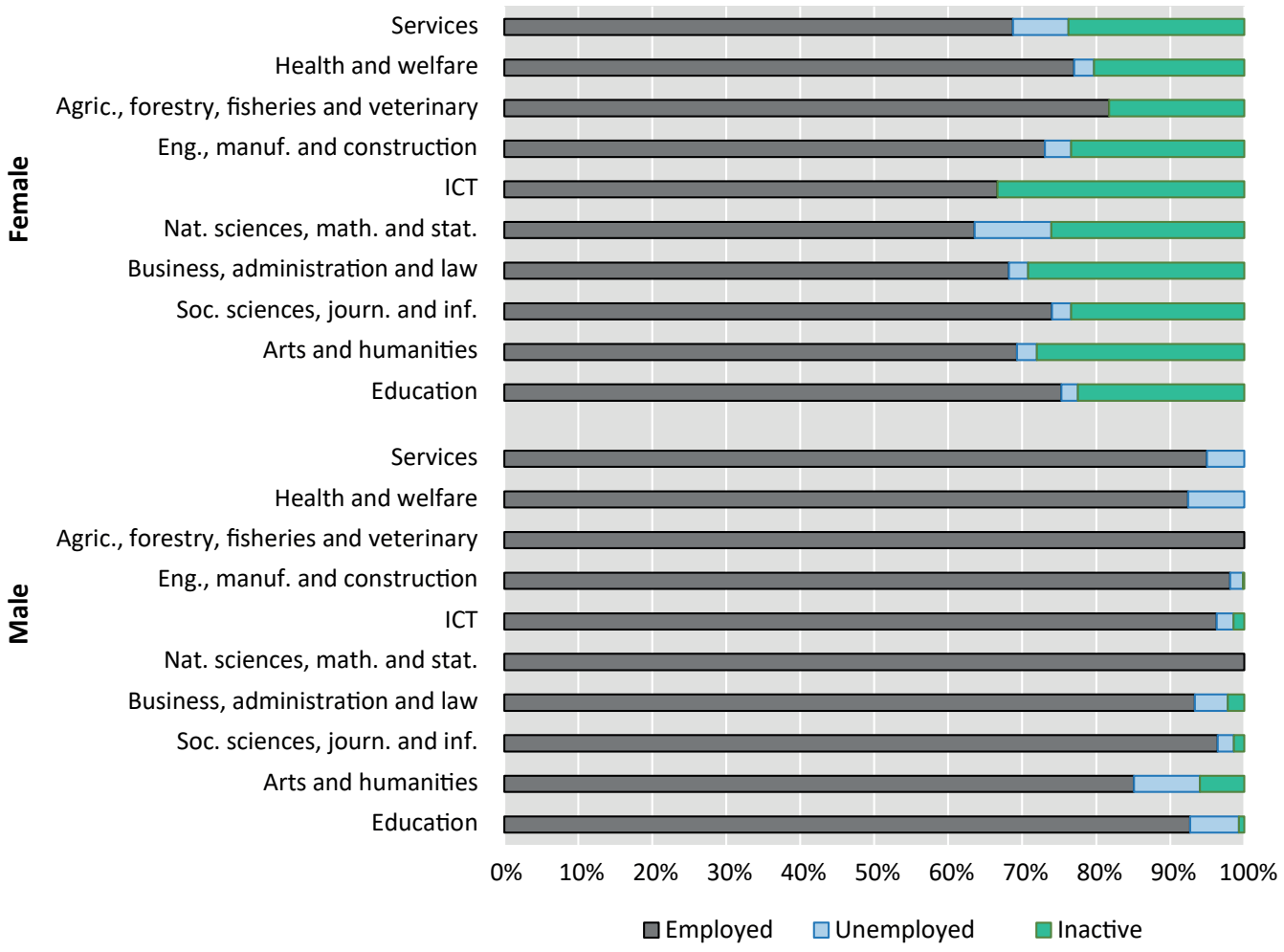
cator. The evolution of overeducation is composition-adjusted to avoid mechanical change in the means due to shifts in demographic composition. As shown in Figure 1, overeducation has not changed much in the observed period. The only statistically significant difference[3] between 2014 and 2019 means was recorded in the case of overeducation 4 indicator (4 pp.). From this point onwards, only the 2019 LFS wave will be considered in the analysis.

The incidence of overeducation differs significantly among study fields (Figure 2). This is true regardless of the chosen educational benchmark. However, study fields with high overeducation incidence reported in one indicator tend to have high incidence also in other considered indicators and vice-versa. The highest levels of overeducation are observed in Agriculture, forestry, fisheries; and Services. For example, the most conservative estimate (Overeducation 2) considers 41% of workers from Services as overeducated while according to modal approach (Overeducation 3) as much as 81% of workers from the same study field are overeducated. On the contrary and regardless of the considered indicator, workers from study fields Information and communication technologies (ICT) and Education experience lowest incidence of overeducation. The lowest estimate (Overeducation 2) considers only more than 3% of ICT graduates as overeducated while the highest (Overeducation 3) considers almost 32% ICT graduates as overeducated.

Not only the incidence of overeducation differs among study fields, but there are also significant differences in employment rates among them. Furthermore, there are also gender differences in employment rates within and between study fields. Figure 3 depicts the differences in the employment, unemployment, and inactivity rates of individuals depending on their field

3 At 5% level of significance.

Figure 3
Distribution of employment status among genders and study fields



Source: calculations based on LFS (2019)

of study and gender. In general, men tend to have markedly higher employment rate in all study fields. Although the lowest employment rate (85%) among males experience graduates in Arts and Humanities, it is still slightly more than the highest employment rate of female graduates in Agriculture (82%) or Health and Welfare (77%) which is the highest share among all study fields for female graduates. Nevertheless, there aren't many differences in unemployment incidence among genders with 3.3% for men and 3.1% for women. However, the inactivity is almost exclusively women's issue with only 1% inactive individuals among men and 24% inactivity rate among women. In particular, almost all of the male graduates in Natural sciences; ICT; and Engineering are employed. However, it is important to note that the large majority of inactive individuals declare that they are looking after children or incapacitated adults (93%) and only a small proportion of them declare the desire for work at the time of the survey (5%).

Regression analysis

Estimation results reported in Table 2 show that graduates in ICT experience lower probability of overeducation than graduates in the reference field (Business, Administration and Law). This difference is statistically significant at 1% level of significance irrespective of the selected dependent variable. The estimated risk of overeducation is lower also in the case of Education; Arts and Humanities which is suggested by negative and statistically significant estimates at 1% level significance level in at least four out of eight model specifi-

cation. Values shown in Table 2 are average marginal effects, therefore estimated marginal effect of field Education in Model 1 suggests that graduates in this field are on average about 20 percentage points less likely to become overeducated than graduates in Business, Administration and Law. Field Health and Welfare also shows lower overeducation risk than the reference field at least 10% level of significance in three out of eight model specifications. However, it also shows higher overeducation risk than the reference field when the second overeducation indicator is considered. On the contrary, graduates in Services suffer from a higher risk of overeducation compared to the reference category. This is suggested by all estimated model specifications. Workers from Natural Sciences, Mathematics and Statistics do not show statistically significant differences in risk of overeducation compared to the reference field.

As mentioned, there are several possible channels through which field of study may affect the risk of overeducation. In this part, we consider labour market conditions, or more precisely, labour market shortages as a determinant of overeducation risk. In Slovakia, there is a strong belief of employers and the general public that the allocation of high-skill labour according to their field of study is not optimal. Moreover, this view is also supported by some evidence (McGowan and Andrews 2015a). To capture the effect of labour market situation, we employ a measurement of labour shortage based on job vacancy data (see Model and variable description). To capture the effect of labour shortages, the job vacancy rate indicator is included in the estimation of the models 2, 4, 6, and 8. The estimated results suggest a strong negative relationship between labour shortages and the probability of overeducation. According to the results of model 2, job vacancy rate higher by 1 standard deviation is associated with a 7 percentage points lower risk of overeducation. These results confirm the important role of the labour market situation in determining the risk of overeducation. Moreover, the inclusion of the job vacancy rate index allows separating the effects of labour shortages that are otherwise presumably captured by field of study dummies. After introducing the job vacancy rate to the model, the coefficients change markedly in the case of fields Agriculture, Forestry, Fisheries, and Veterinary; and Health and Welfare. In the case of Agriculture, the positive effect on overeducation risk decreases in all specifications and become insignificant after controlling for labour market shortages. In the case of the Health and Welfare, the estimated negative effect is reduced after labour shortage index inclusion and becomes insignificant in three out of four overeducation specification. Another interesting finding is that overeducation risk varies across regions, while graduates with residence outside of Bratislava region suffer from higher overeducation risk. This is most likely due to the distribution of jobs in the country since high-skill jobs are concentrated to a large extent in the Bratislava region. More than 55% of jobs in the Bratislava region have high-skill requirements (ISCO 1 - 3) which is markedly higher than the national average (34%)[4].

In terms of education level and years of education, tertiary education graduates are not a homogenous group. Therefore, we distinguish between graduates of Bachelor and Master programmes in the model specification. Graduates of Master programmes exhibit a lower risk of overeducation in three out of four model settings. This finding may support frequent doubts regarding the signalling power of a Bachelor´s programmes in Slovakia. This is suggested by the second-lowest share of Bachelor graduates in the working-age population (Machlica *et al.* 2017) who rather pursue their studies in Master programmes. Only 15.5% of individuals in the restricted sample are Bachelor graduates. However, seemingly clear results are disrupted by contradictory results obtained from the first model specification. According to the first estimated model, graduates holding a Master´s degree are about 12 percentage points more likely to become overeducated. However, these contradictory findings are most likely based on a different indicator specification[5].

4 Calculations based on Slovak LFS (2019).

5 In the case of Overeducation 1,64% of jobs have lower educational requirements than Bachelor´s degree and 83% of jobs lower than Master´s degree. In the case of Overeducation 2,61% of jobs have lower requirements than Bachelor´s degree and 66% lower than Master´s degree. Indicators Overeducation 3 and 4 do not distinguish between Master´s and Bachelor´s requirement, but only consider requirements of tertiary education. 84% of jobs have lower requirements than tertiary education in the case of Overeducation 3 and 67% in the case of Overeducation 4.

Table 2
Average marginal effects of overeducation determinants

Variables	(1) Overed 1	(2) Overed 1	(3) Overed 2	(4) Overed 2	(5) Overed 3	(6) Overed 3	(7) Overed 4	(8) Overed 4
Field of study (Ref. = Bus., admin., law)								
Education	-0.20***	-0.19***	-0.00	-0.00	-0.23***	-0.23***	-0.02	-0.02
	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)
Arts and humanities	-0.19***	-0.19***	-0.01	-0.01	-0.22***	-0.21***	-0.06	-0.05
	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)
Soc. sciences, journ. and information	-0.05	-0.06	-0.06	-0.06*	-0.04	-0.04	0.00	0.00
	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)
Natural sciences, math. and stat.	0.01	0.02	0.08	0.08	-0.00	0.01	0.05	0.06
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
ICT	-0.32***	-0.31***	-0.21***	-0.21***	-0.36***	-0.35***	-0.24***	-0.23***
	(0.09)	(0.09)	(0.03)	(0.03)	(0.08)	(0.08)	(0.05)	(0.05)
Eng., manuf. and construction	-0.00	-0.02	-0.06	-0.07*	0.02	-0.01	-0.06	-0.07
	(0.06)	(0.06)	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)	(0.05)
Agric., for., fish. and veterinary	0.13	0.03	0.08	0.01	0.13*	0.01	0.27***	0.17
	(0.08)	(0.10)	(0.08)	(0.09)	(0.07)	(0.10)	(0.08)	(0.10)
Health and welfare	-0.15***	-0.01	0.05	0.16*	-0.10*	0.06	-0.10**	-0.01
	(0.06)	(0.08)	(0.05)	(0.08)	(0.05)	(0.08)	(0.05)	(0.08)
Services	0.16***	0.15**	0.18***	0.16**	0.18***	0.16***	0.15**	0.13**
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Female	0.02	0.02	-0.00	-0.00	0.02	0.03	0.03	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Degree (Ref. = Bach. degree)								
Master degree	0.12***	0.12***	-0.13***	-0.13***	-0.22***	-0.22***	-0.24***	-0.24***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Age group (Ref. = 35 - 39)								
20 - 24	-0.09	-0.09	-0.08*	-0.08	0.04	0.05	-0.06	-0.05
	(0.07)	(0.07)	(0.05)	(0.05)	(0.07)	(0.07)	(0.05)	(0.06)
25 - 29	0.04	0.04	-0.01	-0.01	0.03	0.03	0.03	0.03
	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
30 - 34	0.07**	0.07**	0.03	0.03	0.07**	0.07**	0.05*	0.05*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Tenure/10	-0.04	-0.04	-0.10***	-0.09***	-0.04	-0.04	-0.10***	-0.10***
	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Part-time (Ref. = Full-time job)	0.15	0.16	0.10	0.10	0.14	0.15	0.04	0.04
	(0.11)	(0.10)	(0.09)	(0.09)	(0.11)	(0.10)	(0.08)	(0.08)
Presence of unempl. or inactive pers.	0.08***	0.09***	0.03	0.03	0.11***	0.11***	0.07**	0.07**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Region (Ref. = Bratislava reg.)								
Western Slovakia	0.13***		0.08***		0.13***		0.07**	
	(0.04)		(0.03)		(0.04)		(0.03)	
Central Slovakia	0.04		0.02		0.07		0.01	
	(0.04)		(0.03)		(0.04)		(0.04)	
Eastern Slovakia	0.09**		0.05		0.07*		0.04	
	(0.04)		(0.03)		(0.04)		(0.03)	
Job vacancy rate (standardized)		-0.07**		-0.05*		-0.08***		-0.05**
		(0.03)		(0.03)		(0.03)		(0.03)
Observations	4,533	4,533	4,533	4,533	4,533	4,533	4,533	4,533

Notes: estimates are from the probit model with sample selection. Models 1 - 8 represent different specification of the dependent variable where the dependent variable in the first two models is Overeducation 1. Values shown are average marginal effects, corresponding to the average impact of a change in the explanatory variable on the probability of overeducation. Clustered standard errors are shown in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Source: calculations based on LFS (2019)

According to estimated models, there are no significant differences in overeducation risk between men and women. Graduates with longer job tenure exhibit lower risk of overeducation in all considered models while the effect is statistically significant at least at 1% level of significance in two out of four model settings. The model estimates suggest that other things being equal, increase in job tenure by ten years will on average result in 4 - 10 percentage point lower probability of overeducation depending on chosen overeducation measurement. Decreasing probability of overeducation with increasing job tenure is in line with career mobility theory (Sicherman and Galor 1990) which ascribes the positive effect of job tenure on career mobility within a firm.

Gender-specific field of study effects

Empirical research admits that the effects of study fields may differ among genders (Rossen *et al.* 2019). Therefore, in our estimation, we allow for the gender-specific field of study effects. Estimated results confirm that the difference in risk of overeducation between males and females varies in some study fields. Women are much more likely to become overeducated in Social Sciences, Journalism and Information. The robustness of estimated results is supported by positive and strongly significant estimates regardless of the considered educational benchmark. On the other hand, females are less likely to become overeducated in Education. However, the evidence of gender-specific effects in other fields is more or less ambiguous. However, only little is known about gender-specific

Table 3
Conditional marginal effects of being female on overeducation risk

Variables	(1) Overeducation 1	(2) Overeducation 2	(3) Overeducation 3	(4) Overeducation 4
Education	-0.287***	-0.209**	-0.253**	-0.232**
	(0.107)	(0.0916)	(0.104)	(0.0920)
Arts and humanities	0.0915	-0.0139	0.0706	0.0525
	(0.123)	(0.106)	(0.123)	(0.106)
Soc. sciences, journ. and inf.	0.259***	0.188***	0.263***	0.205***
	(0.0706)	(0.0478)	(0.0688)	(0.0606)
Business, administration and law	0.0623	0.0441	0.0562	0.120**
	(0.0645)	(0.0550)	(0.0628)	(0.0586)
Nat. sciences, math. and stat.	0.160	0.0689	0.134	0.0107
	(0.122)	(0.118)	(0.125)	(0.121)
ICT	0.0278	-0.0351*	-0.0804	0.0391
	(0.161)	(0.0203)	(0.143)	(0.0878)
Eng., manuf. and construction	0.0404	0.0155	0.0262	0.120
	(0.0882)	(0.0677)	(0.0856)	(0.0781)
Agric., for., fish. and veterinary	-0.0749	-0.0856	-0.103	-0.0379
	(0.135)	(0.157)	(0.130)	(0.153)
Health and welfare	-0.0504	-0.104	0.0386	-0.115
	(0.0915)	(0.0889)	(0.0872)	(0.0814)
Services	-0.0168	0.0312	-0.0371	0.0224
	(0.103)	(0.114)	(0.0891)	(0.113)

Notes: estimates are from the probit model with sample selection. Models 1 - 4 represent different specification of the dependent variable where the dependent variable in the first model is Overeducation 1. Values shown are average marginal effects, corresponding to the average impact of a change in the explanatory variable on the probability of overeducation. Clustered standard errors are shown in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Source: calculations based on LFS (2019)

preferences in terms of field of study choice in Slovakia. Although the proportion of women in the restricted sample is 61%, females are markedly overrepresented in some fields (Education; and Health and welfare) and underrepresented in others (ICT; Engineering, manufacturing and construction). Notably, overeducation risk for females is lower in Education where women are overrepresented suggesting the presence of different gender-specific preferences already at the time of field of study decision making. Moreover, even men who graduated in Education may more often prefer overeducated jobs due to low attractivity of the teaching profession. According to a survey focused on the attractiveness of the teaching profession in Slovakia (Perignáthová 2019), a low salary in teaching professions discourages men more often than women from pursuing a career in Education.

4. Conclusions and policy implications

The main objective of the study was to explore the relationship between the study field and the risk of overeducation. Empirical results provided in the analysis have confirmed the hypothesis that the probability of overeducation differs markedly among study fields. Graduates in ICT; Arts and Humanities; Education experience significantly lower probability of overeducation than graduates in other fields. On the contrary graduates in Services; Agriculture, Forestry, Fisheries and Veterinary exhibit higher risk of overeducation compared to other study fields. The robustness of the results is supported by the sensitivity analysis considering four different overeducation measurements. However, there are several channels through which field of study may have an impact on the overeducation risk. From the policy perspective, it is important to identify those channels since different channels imply different policy tools to cope with potential negative effects. For these reasons, we included the indicator of labour shortage in the analysis to explore the role of labour market conditions as one of the possible channels through which field of study may affect the probability of overeducation. Indeed, the results show overeducation risk tends to be lower in fields with a higher labour shortage and vice-versa. Nevertheless, these findings suggest that policymakers should be concerned with the allocation mechanism of high-skilled labour. According to Ghignoni and Verashchagina (2014) policy initiatives targeted on reducing educational mismatch are expected to be more effective when they concern supply-side factors. Improving the process of students' allocation into study fields in line with current and future labour market needs may improve the matching process. Quintini (2011) emphasize the role of career counselling and guidance to better adapt to labour market needs. According to Machlica *et al.* (2017), the availability of relevant information regarding graduates´ labour market outcomes is low in Slovakia compared to good practice in OECD countries. Improving the availability of relevant information combined with effective career guidance may therefore improve the allocation process. Matching of available tertiary education programmes to labour market needs may be also relevant. However, the evidence of the efficiency of such policies is insufficient since these kinds of policies are difficult to measure (McGowan and Andrews 2015a).

Another interesting finding that was not the original intention of the study is that overeducation risk varies among regions, with the least risk appearing in the Bratislava region. This is presumably due to the higher concentration of high-skill jobs in the region. Regional differences in overeducation risk suggest the importance of policies that promote residential mobility which is quite low in Slovakia compared to OECD countries (Andrews *et al.* 2011). Labour mobility is strongly connected to housing policies. For example, according to McGowan and Andrews (2015a) Slovak republic has the highest costs of obtaining a building permit among OECD countries while decreasing these costs to the median OECD level would be associated with a 2 percentage points decrease in skill mismatch. Furthermore, since high-skill jobs are mainly concentrated in the Bratislava region, policies that promote high-skill job creation also in less developed regions may reduce the misallocation of high-skilled labour. In general, enhancing technological level through improving the efficiency of Research and development is expected to improve job-education matches (Ghignoni and Verashchagina 2014).

Besides, we showed that some fields of study effects are gender-specific. In particular, women in Social Sciences, Journalism and Information exhibit a higher risk of overeducation than male graduates in the same study field. On the contrary, females are less likely to become overeducated in Education. Notably,

women in Education are also markedly overrepresented suggesting the presence of different gender-specific preferences already at the time of field of study decision making. Teacher´s survey also suggests that salaries in the teaching profession are less attractive for men who graduated in Education compared to women (Perignáthová 2019). Male graduates in Education therefore may prefer overeducated jobs to lower earnings in the field. The evidence of the presence of gender-specific effects in other study fields is however ambiguous.

We are well aware of the limitations of the study findings. Although the empirical results provide evidence that there is a strong relationship between study fields and overeducation risk, it reflects correlations between study fields and overeducation, rather than causal effects. In particular, we are not able to observe an individual's motivation and ability which may have an important impact not only on labour market success but also on the selection of students into study fields. Indeed, on aggregate skill measurements from Programme for the International Assessment of Adult Competencies (PIAAC) show significant differences in measured skills among graduates from different study fields suggesting non-random allocation of students into study fields. Furthermore, there is some evidence that the field of study decision may be determined by an individual's ability. Martinák and Varsík (2020) showed that the average score of students of Slovak universities obtained from final upper-secondary graduation tests markedly differs among tertiary education study fields. The least academically successful upper-secondary graduates that continue their studies at Slovak higher education institutions often pursue further studies in Services and Engineering, Manufacturing and Construction, whereas the most successful continue in fields like ICT and Mathematics; Health and Welfare; and Natural sciences. Moreover, also the quality of higher education institutions may differ among the study fields. Since several studies (McGuiness 2003; Ordine and Rose 2009) emphasize the role of educational quality, more need to be known regarding the relationship between the quality of higher education institutions and educational mismatches in Slovakia. Moreover, recent evidence shows the difference between overeducation and overskilling and also in their impact on productivity (McGowan and Andrews 2015a). Therefore, to allow better inferences related to educational and skill mismatches, enhancing data-gathering initiatives in the area of skills and skills mismatches is desirable.

Finally, when discussing the issue of overeducation it is important to mention that overeducation may not always be perceived only in a negative way. Firstly, in line with career mobility theory, the provided evidence suggests that a part of overeducation is only temporary. The presented estimates suggest a decreasing probability of overeducation with higher job tenure. Another important aspect of overeducation may be associated with positive externalities of education. Moretti (2004) estimated that the increase in the supply of college graduates creates positive spillovers in terms of increased wages of college as well as non-college graduates. There are also well-known non-market benefits of higher educational attainment from which we can mention better health prospects, social capital, higher life satisfaction, and lower crime rates.

Appendix 1: Estimates of the Probit model with sample selection

Variables	Mod. 1a Outcome	Mod. 1a Select.	Mod. 1b Outcome	Mod. 1b Select.	Mod. 2a Outcome	Mod. 2a Select.	Mod. 2b Outcome	Mod. 2b Select.	Mod. 3a Outcome	Mod. 3a Select.	Mod. 3b Outcome	Mod. 3b Select.	Mod. 4a Outcome	Mod. 4a Select.	Mod. 4b Outcome	Mod. 4b Select.
Field of study (Ref. = Bussiness, admin. and law)																
Education	-0.02	0.14	-0.03	0.13	0.43	0.14	0.42	0.12	-0.17	0.15	-0.18	0.13	0.50*	0.13	0.49*	0.12
	(0.29)	(0.33)	(0.29)	(0.32)	(0.28)	(0.32)	(0.28)	(0.32)	(0.28)	(0.33)	(0.28)	(0.32)	(0.27)	(0.33)	(0.27)	(0.32)
Arts and humanities	-0.56*	-0.34	-0.50*	-0.44	0.05	-0.33	0.09	-0.44	-0.62**	-0.33	-0.55*	-0.44	-0.10	-0.35	-0.06	-0.45
	(0.29)	(0.34)	(0.29)	(0.35)	(0.32)	(0.35)	(0.31)	(0.35)	(0.29)	(0.34)	(0.29)	(0.35)	(0.31)	(0.34)	(0.31)	(0.35)
Soc. sciences, journ. and inf.	-0.42**	0.34	-0.43**	0.35	-0.75***	0.33	-0.75***	0.35	-0.41**	0.34	-0.41**	0.35	-0.19	0.32	-0.19	0.34
	(0.19)	(0.27)	(0.19)	(0.27)	(0.26)	(0.27)	(0.26)	(0.27)	(0.19)	(0.27)	(0.19)	(0.27)	(0.21)	(0.27)	(0.21)	(0.27)
Natural sciences, math. and stat.	-0.11	4.72***	-0.08	4.90***	0.22	5.31***	0.23	5.29***	-0.11	4.63***	-0.06	4.71***	0.33	4.56***	0.34	4.54***
	(0.24)	(0.24)	(0.24)	(0.25)	(0.26)	(0.24)	(0.27)	(0.49)	(0.24)	(0.23)	(0.24)	(0.26)	(0.26)	(0.23)	(0.26)	(0.25)
ICT	-0.80***	0.41	-0.75***	0.35	-1.05***	0.41	-1.00***	0.35	-0.75***	0.41	-0.70***	0.35	-0.95***	0.39	-0.91***	0.34
	(0.21)	(0.35)	(0.21)	(0.35)	(0.30)	(0.35)	(0.30)	(0.35)	(0.20)	(0.35)	(0.20)	(0.35)	(0.25)	(0.35)	(0.25)	(0.35)
Eng., manuf. and construction	0.06	0.63**	-0.00	0.69**	-0.15	0.63**	-0.20	0.69**	0.12	0.63**	0.05	0.69**	-0.17	0.62**	-0.22	0.68**
	(0.16)	(0.27)	(0.16)	(0.27)	(0.18)	(0.27)	(0.18)	(0.27)	(0.16)	(0.27)	(0.16)	(0.27)	(0.17)	(0.27)	(0.18)	(0.27)
Agric., for., fish. and veterinary	0.60*	4.65***	0.33	5.43***	0.48	5.23***	0.25	5.83***	0.67*	4.56***	0.35	5.24***	1.01***	4.49***	0.74**	5.07***
	(0.35)	(0.19)	(0.37)	(0.30)	(0.32)	(0.19)	(0.36)	(0.29)	(0.36)	(0.19)	(0.39)	(0.29)	(0.32)	(0.21)	(0.36)	(0.31)
Health and welfare	-0.23	0.03	0.11	-0.64*	0.41	0.03	0.69**	-0.66*	-0.25	0.03	0.17	-0.64*	0.06	0.03	0.37	-0.64*
	(0.25)	(0.32)	(0.30)	(0.38)	(0.26)	(0.32)	(0.32)	(0.38)	(0.24)	(0.32)	(0.29)	(0.38)	(0.27)	(0.32)	(0.32)	(0.38)
Services	0.57**	0.32	0.50**	0.34	0.57**	0.32	0.50**	0.34	0.70**	0.31	0.65**	0.33	0.59**	0.31	0.53**	0.33
	(0.24)	(0.35)	(0.24)	(0.34)	(0.24)	(0.35)	(0.24)	(0.34)	(0.28)	(0.35)	(0.28)	(0.34)	(0.24)	(0.35)	(0.24)	(0.35)
Female	0.09	0.30	0.09	0.30	0.06	0.30	0.06	0.30	0.07	0.31	0.07	0.30	0.25	0.29	0.25	0.29
	(0.17)	(0.22)	(0.18)	(0.22)	(0.19)	(0.22)	(0.19)	(0.22)	(0.18)	(0.22)	(0.18)	(0.22)	(0.18)	(0.22)	(0.18)	(0.22)
Interaction of female with:																
Education	-0.98***	-0.07	-0.97***	-0.06	-0.90***	-0.06	-0.90***	-0.05	-0.90***	-0.08	-0.88***	-0.06	-1.13***	-0.07	-1.12***	-0.06
	(0.33)	(0.37)	(0.33)	(0.36)	(0.33)	(0.36)	(0.33)	(0.36)	(0.33)	(0.37)	(0.33)	(0.36)	(0.32)	(0.37)	(0.32)	(0.36)
Arts and humanities	0.09	0.34	0.03	0.39	-0.18	0.32	-0.23	0.38	0.04	0.34	-0.02	0.39	-0.17	0.34	-0.20	0.40
	(0.38)	(0.41)	(0.37)	(0.41)	(0.42)	(0.41)	(0.41)	(0.41)	(0.39)	(0.41)	(0.38)	(0.41)	(0.41)	(0.41)	(0.40)	(0.41)
Soc. sciences, journ. and inf.	0.53**	-0.29	0.53**	-0.30	0.80**	-0.29	0.78**	-0.30	0.57**	-0.29	0.56**	-0.30	0.33	-0.28	0.33	-0.29
	(0.25)	(0.32)	(0.25)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.26)	(0.32)	(0.26)	(0.32)	(0.27)	(0.32)	(0.27)	(0.32)
Natural sciences, math. and stat.	0.26	-5.32***	0.24	-5.55***	0.06	-5.91***	0.05	-5.94***	0.20	-5.23***	0.19	-5.36***	-0.32	-5.16***	-0.31	-5.18***
	(0.37)	(0.35)	(0.37)	(0.36)	(0.39)	(0.35)	(0.40)	(0.56)	(0.38)	(0.35)	(0.39)	(0.37)	(0.39)	(0.35)	(0.39)	(0.36)
ICT	-0.15	-0.25	-0.21	-0.17	-4.61***	-0.25	-4.61***	-0.19	-0.51	-0.24	-0.58	-0.17	-0.17	-0.24	-0.21	-0.17
	(0.53)	(0.57)	(0.55)	(0.58)	(0.38)	(0.56)	(0.36)	(0.58)	(0.52)	(0.57)	(0.54)	(0.59)	(0.58)	(0.57)	(0.60)	(0.58)
Eng., manuf. and construction	-0.11	-0.73**	-0.11	-0.70*	-0.14	-0.73**	-0.13	-0.70*	-0.14	-0.73**	-0.14	-0.70*	-0.02	-0.72*	-0.01	-0.69*
	(0.29)	(0.37)	(0.29)	(0.37)	(0.32)	(0.37)	(0.32)	(0.37)	(0.29)	(0.37)	(0.29)	(0.37)	(0.30)	(0.37)	(0.30)	(0.37)
Agric., for., fish. and veterinary	-0.42	-4.13***	-0.46	-4.35***	-0.41	-4.69***	-0.43	-4.74***	-0.53	-4.04***	-0.56	-4.17***	-0.48	-3.97***	-0.49	-4.00***
	(0.47)	(0.35)	(0.47)	(0.36)	(0.48)	(0.35)	(0.48)	(0.36)	(0.48)	(0.34)	(0.47)	(0.36)	(0.44)	(0.35)	(0.44)	(0.37)
Health and welfare	-0.30	-0.03	-0.27	-0.04	-0.45	-0.02	-0.42	-0.03	-0.05	-0.03	-0.02	-0.04	-0.76**	-0.03	-0.74**	-0.04
	(0.30)	(0.36)	(0.30)	(0.36)	(0.31)	(0.36)	(0.31)	(0.36)	(0.29)	(0.36)	(0.29)	(0.36)	(0.32)	(0.36)	(0.32)	(0.36)
Services	-0.23	-0.55	-0.19	-0.55	-0.08	-0.56	-0.04	-0.56	-0.31	-0.54	-0.28	-0.54	-0.31	-0.55	-0.27	-0.55
	(0.36)	(0.42)	(0.36)	(0.42)	(0.35)	(0.42)	(0.35)	(0.42)	(0.39)	(0.42)	(0.39)	(0.42)	(0.35)	(0.42)	(0.36)	(0.42)
Master degree	0.34***	0.04	0.33***	0.03	-0.43***	0.04	-0.44***	0.03	-0.65***	0.04	-0.66***	0.03	-0.74***	0.04	-0.74***	0.03
	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)

Variables	Mod. 1a Outcome	Mod. 1a Select.	Mod. 1b Outcome	Mod. 1b Select.	Mod. 2a Outcome	Mod. 2a Select.	Mod. 2b Outcome	Mod. 2b Select.	Mod. 3a Outcome	Mod. 3a Select.	Mod. 3b Outcome	Mod. 3b Select.	Mod. 4a Outcome	Mod. 4a Select.	Mod. 4b Outcome	Mod. 4b Select.
Age group (Ref. = 35 - 39)																
20 - 24	-0.26	-0.68***	-0.24	-0.70***	-0.36	-0.68***	-0.33	-0.70***	0.12	-0.68***	0.15	-0.70***	-0.22	-0.68***	-0.20	-0.70***
	(0.20)	(0.19)	(0.20)	(0.19)	(0.24)	(0.19)	(0.25)	(0.19)	(0.20)	(0.19)	(0.20)	(0.19)	(0.23)	(0.19)	(0.23)	(0.19)
25 - 29	0.11	0.02	0.11	0.01	-0.04	0.02	-0.03	0.01	0.09	0.02	0.09	0.01	0.12	0.02	0.12	0.01
	(0.11)	(0.12)	(0.11)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)
30 - 34	0.20**	0.10	0.20**	0.09	0.10	0.10	0.11	0.09	0.21**	0.10	0.21**	0.09	0.17*	0.09	0.17*	0.08
	(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)
Has children less 3 yrs. old		-0.10		-0.06		-0.10		-0.07		-0.10		-0.06		-0.11		-0.07
		(0.29)		(0.29)		(0.29)		(0.29)		(0.29)		(0.29)		(0.29)		(0.29)
Female * Has child. less 3 yrs. old		-1.92***		-1.96***		-1.92***		-1.95***		-1.91***		-1.95***		-1.90***		-1.95***
		(0.31)		(0.31)		(0.31)		(0.31)		(0.31)		(0.31)		(0.31)		(0.31)
Unemployed or inactive person in the household	0.04	-0.26	0.06	-0.26	-0.11	-0.26	-0.10	-0.26	0.10	-0.26	0.11	-0.27	-0.03	-0.26	-0.03	-0.26
	(0.11)	(0.18)	(0.11)	(0.18)	(0.13)	(0.18)	(0.12)	(0.18)	(0.11)	(0.18)	(0.11)	(0.18)	(0.12)	(0.18)	(0.12)	(0.18)
Female * Unemp. or inact. person	0.36**	0.43*	0.36**	0.43*	0.37**	0.43*	0.37**	0.43*	0.42**	0.43*	0.42**	0.43*	0.43**	0.43*	0.44**	0.43*
	(0.17)	(0.25)	(0.17)	(0.25)	(0.19)	(0.25)	(0.19)	(0.25)	(0.17)	(0.25)	(0.17)	(0.25)	(0.17)	(0.25)	(0.17)	(0.25)
Partner in the household		1.13***		1.11***		1.13***		1.11***		1.14***		1.12***		1.12***		1.10***
		(0.22)		(0.22)		(0.23)		(0.23)		(0.23)		(0.22)		(0.23)		(0.22)
Female * Partner in the household		-1.51***		-1.49***		-1.50***		-1.49***		-1.52***		-1.50***		-1.50***		-1.48***
		(0.25)		(0.25)		(0.26)		(0.25)		(0.26)		(0.25)		(0.25)		(0.25)
Region (Ref. = Bratislava region)																
Western Slovakia	0.36***	-0.46***			0.31**	-0.46***			0.36***	-0.46***			0.24**	-0.46***		
	(0.11)	(0.14)			(0.12)	(0.14)			(0.11)	(0.14)			(0.12)	(0.14)		
Central Slovakia	0.12	-0.40***			0.07	-0.40***			0.19	-0.40***			0.03	-0.41***		
	(0.12)	(0.15)			(0.14)	(0.15)			(0.12)	(0.15)			(0.13)	(0.15)		
Eastern Slovakia	0.24**	-0.45***			0.18	-0.45***			0.21*	-0.45***			0.15	-0.45***		
	(0.11)	(0.14)			(0.13)	(0.14)			(0.11)	(0.14)			(0.12)	(0.14)		
Tenure/10	-0.10		-0.11		-0.36***		-0.36***		-0.13		-0.13		-0.35***		-0.35***	
	(0.11)		(0.11)		(0.12)		(0.12)		(0.11)		(0.11)		(0.12)		(0.12)	
Part-time (Ref.= Full-time job)	0.43		0.45		0.33		0.33		0.43		0.44		0.13		0.14	
	(0.31)		(0.31)		(0.28)		(0.28)		(0.34)		(0.33)		(0.27)		(0.27)	
Job vacancy rate (Standardized)			-0.20**	0.34***			-0.17*	0.34***			-0.24***	0.34***			-0.18**	0.34***
			(0.08)	(0.11)			(0.10)	(0.11)			(0.09)	(0.11)			(0.09)	(0.11)
Constant	-0.47**	1.66***	-0.27	1.35***	-0.44**	1.67***	-0.28	1.36***	0.48**	1.66***	0.68***	1.35***	-0.13	1.68***	-0.01	1.37***
	(0.20)	(0.25)	(0.19)	(0.23)	(0.22)	(0.25)	(0.20)	(0.23)	(0.21)	(0.25)	(0.19)	(0.23)	(0.21)	(0.25)	(0.19)	(0.23)
Observations	3,721	4,533	3,721	4,533	3,721	4,533	3,721	4,533	3,721	4,533	3,721	4,533	3,721	4,533	3,721	4,533

Notes: estimates are from the Probit models with sample selection. Models represent different specifications of the dependent variable in outcome model and different independent variables specifications. Consequently, the dependent variable in the outcome equation of Model 1 is a binary variable Overeducation 1. The dependent variable in selection equations is a binary variable where 1 means that an individual is employed and 0 not employed. Clustered standard errors are shown in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Source: calculations based on LFS (2019)

Appendix 2: Technical appendix

Calculation of job vacancy rate

As a proxy for the labour shortage, we calculated the job vacancy rate for each region and groups of occupations corresponding to each study field. The correspondence between ISCO-08 classification and ISCED field of study classification (ISCED-F 2013) was based on Eurostat (2019) and Wolbers (2013) methodology. However, the relationship between study fields and corresponding occupations is many-to-many rather than exclusive. Thus, one study field may be appropriate for several professions but also various professions are appropriate for various study fields. Therefore, in terms of suitable professions, some study fields more or less overlap. This is mostly the case of field pairs ICT - Natural Sciences, Mathematics and Statistics; and Social sciences, journalism and information - Business, administration and law whose corresponding occupations completely overlap. Consequently, also the values of job vacancy rate index calculated for these study fields coincide. In line with Eurostat methodology[6], the job vacancy rate was calculated as a comparison of the number of job vacancies published in 2019 with the average employment in the same year calculated using LFS data:

$$JVR_{f,r} = \frac{JV_{f,r}}{\left(JV_{f,r} + EMPL_{f,r} \right)}$$

where $JV_{f,r}$ is the total number of published job vacancies corresponding to study field f in region r. $EMPL_{f,r}$ is the average employment in occupations corresponding to study field f in region r. Only the high skilled occupations are considered when calculating the job vacancy rate. High skilled occupations are defined as occupations from ISCO 1 - 3 categories with tertiary education requirements set by employers. Finally, the index is standardized to have nationally a mean of 0 and a standard deviation of 1.

The validity of exclusion restriction

To test the validity of exclusion restriction we follow the strategy from Ghignoni and Verashchagina (2014) who utilized the modified version of Sargan test. The reasonability behind excluding the considered household characteristics from outcome equation assumes that they must be uncorrelated with the error term of the outcome equation, and thus not affect the probability of overeducation conditional on the explanatory variables included. We test this assumption via separate regressions with the residuals of the estimated overeducation equation as a dependent variable and the household characteristics as the repressors. Regressions are estimated separately for each overeducation indicator and gender. Each regression contains only one household variable at the same time. Consequently, for each household variable and gender, we estimate eight separate regressions (4 overeducation indicators * 2 outcome equation specifications). The considered household characteristics are the presence of children aged less 3 years in the household, presence of unemployed or inactive persons in the household, presence of a partner or spouse in the household. The results suggest the credibility of the exclusion restriction in the case of two household variables (presence of children aged less 3 years in the household and the presence of a partner in the same household). For each of them, the R2 varies between 0 - 0.0007. The presence of unemployed or inactive persons in the household is an exception since the R2 in the female regressions achieves 0.003 in some specifications and also the t-test suggests statistically significant coefficient for this variable in four out of eight residual regression estimations. Therefore, we include this household characteristic and its interaction with gender also in the outcome equations.

6 https://bit.ly/3qUGZjp

References

Andrews D., Caldera Sánchez A., Johansson Å. (2011), *Housing Markets and Structural Policies in OECD Countries*, OECD Economics Department Working Papers, n.836, Paris, OECD Publishing <https://bit.ly/3s6TFVU>

Barone C., Ortiz L. (2011), Overeducation among European University Graduates: A comparative analysis of its incidence and the importance of higher education differentiation. *Higher Education*, 61, n.3, pp.325-337 <https://bit.ly/3pSItKE>

Caroleo F.E., Pastore F. (2013), *Overeducation at a Glance: Determinants and Wage Effects of the Educational Mismatch, Looking at the AlmaLaurea Data*. IZA Discussion Papers, n.7788, Bonn, Institute of Labor Economics (IZA) <http://ftp.iza.org/dp7788.pdf >

Croce G., Ghignoni E. (2012), Demand and Supply of Skilled Labour and Overeducation in Europe: A Country-level Analysis. *Comparative Economic Studies*, n.54, pp.413-439 <https://bit.ly/3pRwtci>

Duncan G., Hoffman S. (1981), The Incidence and Wage Effects of Overeducation, *Economics of Education Review*, 1, n.1, pp.75-86

Eurostat (2019), Skills mismatch experimental indicators <https://bit.ly/399Y0iu>

Ghignoni E., Verashchagina A. (2014), Educational Qualifications Mismatch in Europe. Is it demand or Supply Driven? *Journal of Comparative Economics,* 42, n.3, pp.670-692

Hersch J. (1991), Education Match and Job Match, *The Review of Economics and Statistics*, 73, n.1, pp.140-144

Kucel A., Byrne D. (2008), *Are Over-educated People Insiders or Outsiders? A Case of Job Search Methods and Over-education in UK*, ESRI Working Paper, n.258, Dublin, The Economic and Social Research Institute (ESRI) <https://bit.ly/3b7KChn>

Leuven E., Oosterbeek H. (2011), Overeducation and Mismatch in the Labor Market, in Erik Hanushek E., Welch F. (eds.) *Handbook of the Economics of Education*, vol.4. Oxford, Elsevier B.V., pp.283-326

Machlica G., Toman J., Haluš M., Martinák D. (2017), E*nhancing advanced skills to better meet labour market demand in the Slovak Republic*, OECD Economics Department Working Papers, n.1416, Paris, OECD Publishing <https://bit.ly/3niiQBc>

Martinák D., Varsík S. (2020), *Odliv Mozgov: Analýza absolventov stredných škôl študujúcich na zahraničných vysokých školách*, Educational Policy Institute, Ministry of Education, Science, Research and Sport of Slovak Republic <https://bit.ly/3baMCoR>

Mcgowan A. M., Andrews D. (2015a), *Skill Mismatch and Public Policy in OECD Countries*, OECD Economics Department Working Papers, n.1210, Paris, OECD Publishing <https://bit.ly/3hOlqhg>

Mcgowan A. M., Andrews D. (2015b), *Labour Market Mismatch and Labour Productivity: Evidence from PIAAC data.* Economics Department Working Papers, n.1209 Paris, OECD Publishing <https://bit.ly/3bsvR91>

McGuinness S. (2006), Overeducation in the Labour Market, *Journal of Economic Surveys*, 20, n.3, pp.387-418

McGuinness S. (2003), University Quality and the labour market, *Applied Economics*, 35, pp.1943-1955

Montt G. (2015), *The causes and consequences of field-of-study mismatch: An analysis using PIAAC*, OECD Social, Employment and Migration Working Papers, n.167, Paris, OECD Publishing <https://bit.ly/3oiSsIJ>

Moretti E. (2004), Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data, *Journal of Econometrics*, 121, n.1-2, pp.175-212

Nicaise I. (2001), Human capital, reservation wages and job competition: Heckman´s lambda re-interpreted, *Applied Economics,* 33, n.3, pp.309-315

OECD (2019), *OECD Economic Surveys: Slovak Republic 2019*, Paris, OECD Publishing <https://bit.ly/3pRrwjv>

OECD (2017), *OECD Employment Outlook 2017*, Paris, OECD Publishing <https://bit.ly/3neNNWN>

Ordine P., Rose G. (2009), Overeducation and Instructional Quality: A Theoretical Model and Some Facts. *Journal of Human Capital*, 3, n.1, pp.73-105 <DOI: 10.1086/599836>

Ortiz L., Kucel A. (2008), Do Fields of Study Matter for Over-education? The Cases of Spain and Germany, *International Journal of Comparative Sociology*, 49, n.4-5, pp.305-327

Perignáthová M. (2019), *Atraktivita učiteľského povolania*. Educational Policy Institute, Ministry of Education, Science, Research and Sport of Slovak Republic <https://bit.ly/2KUybeb>

Quintini G. (2011), *Right for the Job: Over-Qualified or Under-Skilled?*, OECD Social, Employment and Migration Working Papers, n.120, Paris, OECD Publishing <https://bit.ly/3b9OGxo>

Reimer D., Noelke C., Kucel A. (2008), Labor Market Effects of Field of Study in Comparative Perspective: An Analysis of 22 European Countries. *International Journal of Comparative Sociology*, 49, n.4-5, pp.233-256 <https://bit.ly/3hMOjdB>

Rossen A., Boll C., Wolf A. (2019), Patterns of Overeducation in Europe: The Role of Field of Study, Bonn, *IZA Journal of Labor Policy*, 9, n.1 <https://bit.ly/3pMgT1i>

Sattinger M. (1993), Assignment Models of the Distribution of Earnings. *Journal of Economic Literature*, 31, n.2, pp.831-880 <https://bit.ly/3pPNVxJ>

Sicherman N. (1991), "Overeducation" in the Labor Market, *Journal of Labor Economics*, 9, n.2, pp.101-122

Sicherman N., Galor O. (1990), A Theory of Career Mobility, *Journal of Political Economy*, 98, n.1, pp.169–192

Slovak Statistical Office (2019), DATAcube database <http://datacube.statistics.sk/>

U21 (2020), *Ranking of National Higher Education Systems 2020*, Melbourne Institute of Applied Economic and Social Research, University of Melbourne

Van Praag B., Van De Ven W. (1981), The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection. *Journal of Econometrics*, 17, n.2, pp.229-252 <https://bit.ly/394sC4H>

Wolbers M. (2013), Job Mismatches and their Labour-Market Effects among School-Leavers in Europe, *European Sociological Review*, 19, n.3, pp.249-266 <DOI:10.1093/esr/19.3.249>

Dávid Martinák
david.martinak@savba.sk
Researcher at the Institute of Economic Research, Slovak Academy of Sciences. His research is primarily focused on labour and education economics, educational and labour market mismatches, and supply and demand for skills.