Educational mismatch in the Czech Labour Market[1]

MARTINA MYSÍKOVÁ
National Training Fund, Opletalova 25,
110 00 Praha 1,
Czech Republic

ABSTRACT

Educational mismatch in labour markets is a phenomenon that has been widely analysed, mainly with respect rising concerns about possible oversupply of graduates. Similarly to most European countries, the Czech Republic has experienced a boom in tertiary education in the last decade. The incidence and the determinants of over- and undereducation vary substantially depending on both the approach of the mismatch measurement and the data source applied. Educational mismatch also reflects in wage levels. First, overeducated workers have lower wages and undereducated workers have higher wages compared to workers with the same education whose jobs match their education. Second, overeducated workers earn more and undereducated workers earn less than their co-workers with exactly the required level of education. The effects are qualitatively the same regardless the data source or approach applied but their sizes differ slightly according to the approach and data applied.

Keywords: educational economics, human capital, overeducation, undereducation, wages
JEL classifications: I21, J24, J31

[1] This paper was supported by grant no. P402/12/G130 'The relationships between skills, schooling and labour market outcomes: A longitudinal study' from the Czech Science Foundation. The author would like to thank Jiří Večerník (Institute of Sociology of the Academy of Sciences of the Czech Republic) for his valuable advice.

1. INTRODUCTION

Educational mismatch in the labour markets is a phenomenon that has recently received a lot of attention, mainly due to growing number of university graduates and related concerns about possible oversupply of graduates. Many studies exist on educational mismatch on the U.S. labour market (among others, McGoldrick and Robst, 1996, analysed overeducation probability; Rubb, 2003, discussed the persistence of overeducation). In the European context, most studies are devoted to Western European countries (wage effects of educational mismatch in Germany were analysed by Bauer, 2002; determinants of overeducation in Belgium were examined by Karakaya, Plasman and Rycx, 2007; determinants and wage effects of overeducation in Portugal were analysed by Kiker, Santos and De Oliveira, 1997), while those focused on Eastern Europe are rather scarce.

Education has changed since the beginning of the 1990s in the Czech Republic. First, similarly to other post-communist countries, new graduates shifted away from technical fields towards business fields during the transition from a planned to a market-based economy (Jeong, Kejak and Vinogradov, 2008). Second, similarly to most European countries, the Czech Republic (CR) has experienced boom in tertiary education in the last decade, and the threat of oversupply of graduates resulting in overeducation (or even unemployment) is thus rather topical. So far, this topic has not been thoroughly addressed in the Czech Republic.

Excess supply of persons with higher education is believed to result in them occupying jobs for which they are overqualified. As a consequence, workers with lower education are pushed out and forced to accept also less demanding jobs. With the number of people receiving higher education further increasing, this problem could intensify even more. According to Freeman (1976), if the excess qualified workers are pushed to 'lower-qualified' jobs, returns to education should fall, which would lower investments in education and as a consequence the labour market would ultimately reach equilibrium. Chevalier (2003) argues that this did not happen in the U.S. because returns to education remained high. Gottschalk and Hansen (2003) did not find any supporting evidence that an increasing share of college graduates had been forced to non-college jobs since the mid-1980s to the mid-1990s in the U.S.

Similarly, studies on the UK argue that although the share of tertiary educated graduates has been increasing over the last decades, the returns to education remained stable (Chevalier, 2003; Bevan and Cowling, 2007). Contrary to expectations, growing share of higher-educated and stable returns from education indicate that demand for graduate workers coped well with the increasing supply. Some researchers look for alternative explanations, for instance, heterogeneity in skills of workers with the same education level (for instance, Chevalier, 2003).

Unlike advanced economies, the information about transitional economies is rather scarce. The aim of this study is to describe the situation in the Czech Republic. In this country, returns to education rose rapidly since the end of the communist era in 1989, reflecting the changed principles of remuneration. According to Večerník (2013) returns to education doubled during the early years of economic transition, but since 1996, they remained stable for both men and women. This suggests that since the late 1990s the 'Western' arguments regarding stable returns to education while rising tertiary enrolments might apply for the CR as well. According to

Eurostat database, the share of people with tertiary education on working-age population increased from 10 percent in 2003 to 17 percent in 2012. This growth mainly occurred among young people: 13 percent of 30-34-old had tertiary education in 2003 while their share rose to 26 percent in 2012.

As there is little empirical evidence on the educational mismatch in the Czech labour market, and in order to view this phenomenon as broadly as possible, this study aims to apply various methodological approaches to measure education required in an occupation known from the literature. Measurement of educational mismatch is highly sensitive to the methodology, as described in Section 2. For that reason, this study uses several data sources to check the robustness of the results. Section 3 describes the EU-SILC and PIAAC data used for the purposes of this paper, and comments on their pitfalls. Section 4 analyses the determinants of over- and undereducation, while Section 5 measures the effects of educational mismatch on wages. Section 6 evaluates the results obtained under different methodological approaches and data sources.

APPROACHES TO MEASUREMENT OF REQUIRED EDUCATION

Analyses of educational mismatch typically face one problem: the results tend to differ according to the applied methodology of how to measure the education required in an occupation. Workers are considered adequately educated (matched), if their education just fits the requirements of their job. In all other cases, they are considered overeducated/undereducated if their level of education exceeds/falls short to the requirements of their job.

Three approaches of how to determine an education level required for an occupation appear in the literature. OECD (2007, p. 135) call them normative, statistical, and self-declared approaches. The *normative* approach relates education and job qualifications based on an exogenous definition of education requirements determined by job analysts (used, for example, by Chevalier, 2003; Karakaya et al., 2007). In European practice, OECD methodology (OECD, 2007) which links the education ISCED codes with ISCO classification according to an established scheme is usually applied, for instance by Hernández and Serrano (2012). Studies on the U.S. labour market typically use the Dictionary of Occupational Titles for this purpose (among others, McGoldrick and Robst, 1996).

The *statistical* approach determines the required years of education from realized job matches. Two measures under this approach have been discussed in literature: required years of education have been established as a) a one-standard-deviation range around the national mean of education in an occupation (Verdugo and Verdugo, 1989; Bauer, 2002; Rubb, 2003) and b) a modal value (suggested by Kiker et al., 1997, applied, for instance, by Bauer, 2002; Karakaya et al., 2007). The third approach adopted by other studies is called the *self-declared approach* and it uses workers' answers on questions about the requirements of their jobs (Nieto, 2014; Alba-Ramírez, 1993; McGoldrick and Robst, 1996).

McGoldrick and Robst (1996) used all three approaches to analyse the situation in the U.S. in 1985. Following the normative approach, they classified more than one half of both men and

women as overeducated. On the other hand, according to the statistical approach (one-standard-deviation measure), only 16 percent of male and 9 percent of female workers were overeducated, and, finally, the self-declared approach referred to about 30 percent of overeducated men and women. Chevalier (2003) notes that the choice of approach highly influences the incidence of overeducation, but not its effects on wages.

Another problem is the assumption embedded in all three approaches: the homogeneity of individuals with the same level of education. Some studies control for the unobserved heterogeneity by using panel estimation techniques (for instance, Bauer, 2002). Chevalier (2003) deals with the heterogeneity of UK graduates by creating a group of 'the clever ones' and 'the under-achievers', based on a question on how dis/satisfied the respondents are with the match between their work and their qualifications (assuming that overeducated who are satisfied are only 'apparently' overeducated, while those who are dissatisfied are 'genuinely' overeducated). Unfortunately, none of the available Czech data sources seem to provide a possibility to adequately examine the homogeneity assumption.

All three approaches of how to measure the required education are subject to critics. The normative approach is often criticised for its arbitrary nature, especially if the same established scheme is applied for all countries, as the level of education required for an occupation can differ from one country to another (for example, OECD, 2007, p.135). The statistical approach is also criticised for certain arbitrariness as there is no logic behind the choice of one standard deviation (McGoldrick and Robst, 1996). On the one hand, it has the advantage of being sensitive to technological changes and labour market characteristics (Karakaya et al., 2007), but on the other hand, the reference group of matched workers itself can be overqualified if majority of workers in an occupation are more educated than actually needed. The advantage of the self-declared approach is that it is job-specific; its disadvantage is naturally its subjectivity tending to overestimate own qualification in regard to the job currently performed.

DATA FOR THE CZECH REPUBLIC AND APPLICABILITY OF EXISTING APPROACHES

Three data sources were used for the analyses. First, the national Labour Force Survey (LFS) from the first quarter of 2013 with a total sample of 57,000 respondents, the biggest sample among similar individual/household surveys in the CR. LFS does not collect information on wages and thus cannot be used for my final analyses. Second, I applied the national version of Statistics on Income and Living Conditions (SILC) from 2012 with a total sample of almost 21,000 respondents. The national dataset includes the variable of educational attainment in a more detailed structure than the international dataset (for more details on differences between Czech national and international datasets, see Mysíková, 2011). And third, the national survey of the Programme for the International Assessment of Adult Competencies (PIAAC), conducted in 2011/2012 in the CR with a total sample of 6,000 respondents was applied.

Working age (16-65) full-time employees working 30+ hours per week were selected in all three datasets. For the normative approach, I follow the OECD methodology for the educational mismatch measurement. According to it, low-skilled workers with ISCED 0-2 education are matched in low-skilled occupations ISCO 9 and undereducated otherwise. Intermediate-skilled workers with ISCED 3-4 education are matched at intermediate occupations ISCO 4-8 (overeducated if they work at ISCO 9 occupations and undereducated if they work at ISCO 1-3 occupations). Finally, highly-skilled workers with ISCED 5-6 education are matched at high-skilled occupations ISCO 1-3 and overeducated otherwise (for the detailed scheme, see OECD, 2007, p. 156).

Educational attainment was rescaled into 9 categories (ISCED 1, 2, 3a-b, 3c, 4, 5a – bachelor, 5a - master, 5b, 6), consistent throughout the datasets. For the statistical approach, years of education were derived according to a national standard length required to complete a particular education level. Due to its large sample size, LFS was used to determine the one-standard-deviation range around the average years of education in an occupation and these required years of education were applied in SILC and PIAAC.[2] The statistical approach is based on 1-digit ISCO. This simplification does not alter the results as the statistical approach applied on LFS brought almost the same results regardless the detail of ISCO classification.[3]

The self-declared approach can be applied only in PIAAC. I utilized a question: 'If applying today, what would be the level of education a potential candidate would need to get this job?' The possible answers correspond to the structure of ISCED codes of respondents' educational attainments. The education levels stated by respondents were determined as the required education in the respondents' jobs.

Table 1 shows that the normative approach yields the lowest share of overeducation (6.0-8.3 percent) and the highest share of undereducation (21.9-26.1 percent) and, hence, undereducation is a far more frequent phenomenon than overeducation. The results are quite different if the other approaches are applied. With the statistical approach, overeducation reached about 10 pp higher share (14.2-18.5 percent), yet the share of undereducation was by far the lowest, especially in LFS and SILC (12.4-12.7 percent). The self-declared approach showed a striking one quarter of respondents as overeducated, the share of matched respondents was by far the lowest.

[2] When SILC or PIAAC means and standard deviation values were used to determine the required years of education in an occupation, the share of overeducation was by 1.8 pp higher and the share of undereducation by 1.8 pp lower in SILC. In PIAAC, the share of overeducation was by 0.5 pp lower and the share of undereducation by 2.8 pp lower, compared to using the LFS values in particular data.
The "mode" measure of the statistical approach was examined as well. It resulted in 5-8 pp higher share of overeducation and 2-6 pp higher share of undereducation. The matched group of workers is defined more strictly by the "mode" than the "mean" measure. The one-standard-deviation range around the mean of derived years of education usually does not include only one education level but also adjacent levels, as against the modal value which determines just one educational level which is the required one in an occupation.

[3] The broader the ISCO categories, the higher were the shares of both over- and undereducated; however, the difference was only moderate: The share of overeducated ranged from 14.2% to 14.5% and the share of undereducated ranged from 10.0% to 12.4% resulting from 3-digit and 1-digit ISCO, respectively.

Table 1 Educational mismatch (%)

	Normative			Statistical			Self-declared
	LFS	SILC	PIAAC	LFS	SILC	PIAAC	PIAAC
Total							
Overeducated	6.6	6.0	8.3	14.5	14.2	18.5	25.4
Matched	71.6	67.9	69.5	73.0	73.2	64.9	59.8
Undereducated	21.9	26.1	22.1	12.4	12.7	16.7	14.9
Male							
Overeducated	5.0	4.9	5.9	17.5	17.2	23.8	25.4
Matched	74.8	73.7	75.0	72.1	71.3	62.8	61.5
Undereducated	20.1	21.4	19.1	10.3	11.4	13.4	13.1
Female							
Overeducated	8.6	7.2	11.3	10.7	10.4	21.3	25.3
Matched	67.4	60.9	63.0	74.1	75.4	55.7	57.7
Undereducated	24.1	31.8	25.7	15.2	14.2	23.0	17.0
Youth (16-24)							
Overeducated	7.5	6.8	8.0[a]	21.0	15.3	20.5[a]	25.0[a]
Matched	72.7	70.1	77.0[a]	68.0	73.2	65.2[a]	64.8[a]
Undereducated	19.8	23.0	15.0[a]	11.0	11.4	14.3[a]	10.2[a]
Elderly (55-65)							
Overeducated	5.9	6.6	4.9[a]	10.2	12.3	21.9[a]	22.2[a]
Matched	69.0	64.0	66.8[a]	71.9	69.2	50.1[a]	53.4[a]
Undereducated	25.1	29.3	28.3[a]	17.9	18.5	28.0[a]	24.4[a]

Source: National LFS 1q2013, national SILC (Životní podmínky) 2011, PIAAC. Author's computations.
Notes: Unweighted sample size: LFS – 18,647; SILC – 7,044; PIAAC – 2,344. [a]Sample size of youth and elderly in PIAAC includes less than 340 observations.

While all the approaches concur that undereducation is higher for women than for men, the normative approach establishes higher overeducation for women, as opposed to the statistical approach, which suggested that men are more often overeducated than women. With the exception of PIAAC data, the bottoms parts of Table 1 show that the incidence of overeducation is decreasing with age. All the data sources and approaches confirm the increasing incidence of undereducation with age.

Differences between the normative and the statistical approaches arise for several clear reasons. First, workers with ISCED 6 (PhD degree) always show as overeducated under the statistical approach, while they are matched if they work in ISCO 1-3 occupations under the normative approach. On the other side of the spectrum, workers with ISCED 1-2 (primary school) always turned undereducated in the statistical approach, although they can be matched if they work in ISCO 9 occupation under the normative approach. These facts disfavour the statistical approach as both PhD holders and workers with primary education can

without a doubt find a job with requirements that fit their education in reality. Nevertheless, there are only 0.6 percent of PhD holders and 4.2 percent of primary-educated in my LFS sample and, hence, the bias caused by the statistical approach might be negligible.

Figure 1 depicts the differences between the normative approach and the mean measure of the statistical approach. In total, the normative approach imposes higher educational requirements than the statistical approach on 23.2 percent workers. While considered overeducated under the statistical approach, 11.2 percent of workers are classified as matched under the normative approach. It concerns mainly ISCO 3, 7 and 8 occupations and workers with ISCED 3a-b (63.9 percent) or ISCED 5a – master degree (27.6 percent). Further, 12.0 percent of workers are classified as undereducated under the normative approach while their education and occupation match under the statistical approach. It concerns ISCO 3 workers mainly with ISCED 3a-b education.

Figure 1 Differences in educational mismatch under normative and statistical approach in LFS

				Statistical				
isco 1	0.8%			overeducated	matched	undereducated	isco 9	100.0%
isco 2	3.9%		overeducated	3.3%	3.2%	-		
isco 3	27.9%	Normative	matched	11.2%	57.8%	2.6%	isco 4	57.1%
isco 5	2.2%		undereducated	-	12.0%	9.9%	isco 9	42.9%
isco 6	2.0%							
isco 7	31.3%			isco 1	0.9%			
isco 8	31.9%			isco 3	99.1%			

Source: National LFS 1q2013. Author's computations.

Contrary to that, compared to the normative approach the statistical approach imposes higher educational requirements on 5.8 percent workers. 3.2 percent of workers are matched according to the statistical approach while they are overeducated according to the normative approach. These employees work solely in ISCO 9 occupations and their education is at ISCED 3c. And, finally, 2.6 percent of workers are undereducated according to the statistical approach and matched according to the normative approach. These employees work solely in ISCO 4 or 9 occupations and their education is either ISCED 2 (42.9 percent) or ISCED 3c (57.1 percent).

Normative approach based on OECD scheme might not be suitable for the CR as it results in relatively low overeducation but relatively high undereducation – rather an opposite problem what we would expect at current labour market with increasing tertiary education. In the Czech case, one third of workers with ISCED 3a-b education are employed in jobs at ISCO 3 for which they are supposed to be undereducated according to the normative approach.

As expected from empirics from other countries (among others, McGoldrick and Robst, 1996, for the U.S.), the outcome of analysis of educational mismatch in the CR is to a large extent determined by the approach applied. However, this section shows that also the data source influences the results as there is a difference of up to 4 pp in both over- and undereducation under the same approach but different data source. Next section examines the determinants

of over- and undereducation and demonstrates the dissimilarities between the different approaches and different datasets.

DETERMINANTS OF EDUCATIONAL MISMATCH

Tables 2 and 3 show the impacts of various characteristics on the likelihood of being overeducated and undereducated under the three approaches based on SILC and PIAAC. Multivariate logit regression was applied and, hence, the estimated logit coefficients have to be interpreted relative to the reference group, which is matched respondents in this case. Regressions were estimated separately for men and women when SILC was used but not when much smaller PIAAC survey was applied.

Under the normative approach (Table 2), compared to women, men are less likely to be both over- and undereducated in their jobs, relative to having a matched education. Self-declared approach (Table 3) shows the same results regarding undereducation. However, the statistical approach provides an opposite result for overeducation: the probability of being overeducated rather than matched is higher for men compared to women. Theoretically, from several reasons, women would be expected to be overeducated more often than men, due to vertical segregation, gender differences in time constraints and geographical mobility (women as secondary earners move more often to partners' job locations and are more likely to be constrained in their job search by family responsibilities). However, the results based on the statistical measurement of overeducation contradict the expectations. Literature offers examples of both higher (for instance, Karakaya et al., 2007, for Belgium) and lower (for example, Kiker et al., 1997, for Portugal) probability for women than for men to be overeducated for their jobs.

The more years of education, the higher the probability of overeducation: My results mostly confirm this positive relation (similarly as Kiker et al., 1997, for Portugal), with the only exception being the normative approach based on SILC data, where the negative impact is given by the female subsample. The existing studies often discuss an argument of possible trade-off between education and other forms of human capital, such as experience or on-the-job training. Workers may substitute lack of work experience by overeducation and accept a job that requires less education than they achieved in order to gain more experience, which would later enable them to improve their job level (Rosen, 1972). In this case, overeducation would be a temporary phenomenon.

To the contrary, deficit years of education might be compensated by experience and, undereducation could be thus long-lasting or permanent. The negative relationship between experience and the probability of being overeducated proved to be statistically significant under the normative approach based on SILC data (only for men) and under the statistical approach based on PIAAC data. However, the positive relationship between experience and the probability of being undereducated did not prove in any case, moreover, rather the contrary holds for men based on SILC data.

Table 2 Determinants of overeducation and undereducation: Normative approach

	SILC						PIAAC	
	Total		*Male*		*Female*		Total	
	Over	Under	*Over*	*Under*	*Over*	*Under*	Over	Under
Years of education	-0.037	-0.327***	*0.009*	*-0.367****	*-0.084***	*-0.343****	0.133***	-0.326***
Male	-0.415***	-0.620***	-	-	-	-	-0.699***	-0.713***
Experience	-0.045**	-0.006	*-0.076****	*-0.033***	*-0.012*	*0.014*	0.001	0.011
Experience^2	0.001***	0.000	*0.002****	*0.001***	*0.000*	*-0.000*	-0.000	-0.000
Prague	0.012	0.653***	*0.052*	*0.641****	*-0.078*	*0.577****	-0.430	-0.084
Supervisory	-0.688***	1.274***	*-0.286*	*1.800****	*-1.569****	*0.778****	-0.787***	1.322***
Unlimited contract	-0.650***	0.216*	*-0.268*	*-0.036*	*-0.882****	*0.422***	-0.524**	0.399**
Firm size 50+	0.083	0.082	*0.161*	*0.193**	*-0.040*	*-0.115*	0.733***	0.065
Constant	-0.876**	2.334***	*-2.052****	*2.654****	*-0.304*	*2.075****	-2.225***	1.297**
Pseudo R^2	0.10		*0.10*		*0.11*		0.12	
LR X^2 (34/32)	1004.2		*506.85*		*562.87*		387.53	

Source: National SILC (Životní podmínky) 2011, PIAAC. Author's computations.
Notes: Weighted. *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level. All regressions include nine dummies for industry.

Table 3 Determinants of overeducation and undereducation: Statistical and self-declared approach

	Statistical								Self-declared	
	SILC						PIAAC		PIAAC	
	Total		*Male*		*Female*		Total		Total	
	Over	Under	*Over*	*Under*	*Over*	*Under*	Over	Under	Over	Under
Years of education	0.540***	-1.144***	*0.466****	*-1.140****	*0.749****	*-1.253****	0.622***	-0.965***	0.278***	-0.397***
Male	0.654***	-0.477***	-	-	-	-	1.039***	-0.939***	-0.107	-0.460***
Experience	0.009	-0.039**	*0.026*	*-0.060****	*0.015*	*-0.020*	-0.049*	0.008	0.000	-0.017
Experience^2	-0.000	0.001***	*-0.001*	*0.001****	*-0.001*	*0.001*	0.001	0.000	0.000	0.001*
Prague	-0.628***	0.436***	*-0.310**	*0.702****	*-1.232****	*-0.019*	-0.468*	0.082	-0.309*	-0.231
Supervisory	-1.250***	2.203***	*-1.252****	*2.544****	*-1.070****	*1.921****	-0.750***	2.432***	-0.873***	1.145***
Unlimited contract	-0.130	0.270	*0.002*	*-0.128*	*-0.371*	*0.722****	0.088	-0.016	-0.290*	-0.046
Firm size 50+	-0.328***	0.183*	*-0.174*	*0.073*	*-0.724****	*0.251*	-0.029	-0.426***	0.009	0.129
Constant	-8.958***	10.828***	*-7.835****	*11.096****	*-11.282****	*11.055****	-10.320***	9.050***	-3.626***	2.532***
Pseudo R^2	0.26		*0.22*		*0.35*		0.32		0.11	
LR X^2 (34/32)	2555.55		*1226.42*		*1468.95*		1138.81		429.23	

Source: National SILC (Životní podmínky) 2011, PIAAC. Author's computations.
Notes: Weighted. *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level. All regressions include nine dummies for industry.

Living in the capital city decreases the probability of overeducation over having a matching job, although there is no significant effect under the normative approach. To the contrary, living in Prague increases the likelihood of being undereducated rather than matched only when SILC data are used. The results thus show that large labour market provides more vacancies and workers are thus given better chances to find a fitting job and they are more often able to ap-

ply for jobs with higher educational requirements. With the PIAAC data no such effect has been observed.

Supervisory positions are less likely to be occupied by overeducated and more likely to be occupied by undereducated workers than non-supervisory positions (relative to matched workers), uniformly throughout the approaches and datasets. These findings seem rather surprising, especially for undereducation, since supervisory positions would be expected to be held by adequately educated or overeducated workers. Imagine, for example, that an overeducated worker is promoted to a position with higher educational requirements and, hence, becomes matched. If a matched worker is to be promoted to a position with higher education requirements, for instance due to her/his other qualifications and skills, s/he becomes undereducated. Indeed, the share of undereducated is higher among those at supervisory positions than among those at non-supervisory positions. This suggests that other skills and qualification characteristics play a role in promotion rather than education.

All approaches except the statistical one confirmed that having a job contract for indefinite duration (compared to all other forms of contracts) mostly decreases the probability of being overeducated and increases the probability of being undereducated rather than matched. The models based on SILC data were run separately for men and women and revealed that this in fact holds only for women. No statistically significant impact has been proved for men. These findings are in contrast to those by Karakaya et al. (2007) for Belgium, where permanent job contract was found to be positively related to the probability of being overeducated.

Van der Meer and Wielers (1996) found a positive relation between firm size and overeducation. The explanation is that as large firms have higher costs on monitoring productivity, they employ more educated workers who are granted more autonomy. Karakaya et al. (2007) did not confirm this assumption on Belgian labour market. The results in this study do not provide any consistent findings. While the company size does not have any statistically significant impact in most cases, workers in large firms (with 50 and more employees) have higher probability than those in smaller firms to be overeducated rather than matched under the normative approach based on PIAAC (Table 2), and lower probability under the statistical approach based on SILC data (Table 3).

EDUCATIONAL MISMATCH AND WAGES

Two models estimating the effects of overeducation and undereducation on wages have been commonly applied in the literature: the first one follows Verdugo and Verdugo (1989) and the second one Duncan and Hoffman (1981). I apply both of them using all the three approaches of educational mismatch measurement and both the datasets SILC and PIAAC. While the previous sections showed that both the shares of over- and undereducation and the impacts of various determinants on the probability of being over- and undereducated differ according to the approach and data used, the empirics suggests that results of wage effects are rather consistent.

Verdugo and Verdugo model (1989) can be written as:

$$lnY_i = \alpha_0 + \alpha_1 E_i + \alpha_2 OVER_i + \alpha_3 UNDER_i + X_i\gamma + \varepsilon_i, \quad (1)$$

where lnY_i is the log of gross hourly wages of individual i, E_i denotes actual years of education attained, $OVER_i$ and $UNDER_i$ are dummies for overeducation and undereducation, respectively, X_i is a vector including other explanatory variables with the vector of coefficients γ, and ε_i is an error term.

Verdugo and Verdugo model includes years of education attained, and compares mismatched and matched workers with the same level of education. If productivity and wages were determined by the actual years of education attained, then the required education in a particular job and the fact that an individual is over- or undereducated in that job would be irrelevant to wage determination and the coefficients α_2 and α_3 would be zero. In other words, only actual years of education would be relevant to wage level. If $\alpha_2 = \alpha_3 = 0$, equation (1) would be reduced to a standard human capital Mincerian (1974) wage equation:

$$lnY_i = \alpha_0 + \alpha_1 E_i + X_i\gamma + \varepsilon_i. \quad (2)$$

If, on the other hand, wages are related to the required years of education in a particular job, any additional years of actual education would not be rewarded and would not increase the individual's wage. Such overeducated worker would earn less than a similar but matched worker with the same actual education whose job has higher educational requirements. Hence, coefficient α_2 would be negative. In the same logic, an undereducated worker would earn more than an equally educated worker who works in a different job with lower educational requirements that fit his actual education. The last row of Table 4, which states the results of equation (1) using OLS regression, shows that based on computed F-statistics the null hypothesis $\alpha_2 = \alpha_3 = 0$ can be rejected regardless the approach and data used.

Table 4 Wage models: Verdugo–Verdugo

	Normative				Statistical				Self-declared
	SILC			PIAAC	SILC			PIAAC	PIAAC
	Total	***Male***	***Female***	**Total**	**Total**	***Male***	***Female***	**Total**	**Total**
Years of education	0.071*	*0.056**	*0.088**	0.060*	0.075*	*0.061**	*0.095**	0.069*	0.064*
OVEReducated	-0.164*	*-0.150**	*-0.162**	-0.191*	-0.102*	*-0.082**	*-0.140**	-0.146*	-0.161*
UNDEReducated	0.117*	*0.093**	*0.146**	0.137*	0.080*	*0.078**	*0.103**	0.104*	0.073*
R^2	0.43	*0.35*	*0.48*	0.44	0.42	*0.35*	*0.45*	0.42	0.43
H_0: $\alpha_2 = \alpha_3$	160.12*	*48.10**	*129.70**	467.96*	57.97*	25.18*	41.89*	37.09*	70.70*

Source: National SILC (Životní podmínky) 2011, PIAAC. Author's computations.
Notes: Weighted. Unweighted sample size: SILC – 6,471 (3,511 men, 2,960 women); PIAAC – 2,034 (and 2,022 in self-declared approach).
* statistically significant at the 1% level. All regressions include a dummy for sex, experience and experience squared, a dummy for Prague region, a dummy for supervisory position, a dummy for unlimited job contract, a dummy for large company size and nine dummies for industry. Male and female wage equations include also a dummy for parenthood and a dummy for marriage (as they have an opposite effect on wages of men and women).

Overeducated workers earn 10 percent (statistical approach in SILC) to 19 percent (normative approach in PIAAC) less than workers with the same education who work in jobs that just fit their education. Normative approach provides higher wage penalty than the statistical and

self-declared ones. Regarding the data source, PIAAC shows higher wage penalty than SILC for both normative and statistical approaches. The wage penalty is higher for women than for men. Bauer (2002) showed slightly higher results using German Socio-Economic Panel data and the statistical approach: wage penalty for overeducated men was 10.6 percent and 15.1 percent for overeducated women, compared to 8.2 percent and 14.0 percent, respectively, in this study. As shown by Kiker et al. (1997), the wage penalty was higher for men (9.5 percent) than for women (3.6 percent) in Portugal in 1991.

Undereducated workers receive from 7 percent (self-declared approach in PIAAC) to 14 percent (normative approach in PIAAC) more than workers with the same education in matched jobs. The wage premium is the highest according to the normative approach and the lowest according to the self-declared approach. PIAAC provides higher wage premiums than SILC. The wage premium is higher for women than for men. Considering the statistical approach, Bauer (2002) showed similar wage premium for undereducated men (8 percent) but much lower for women (3 percent) in Germany, and Kiker et al. (1997) about 16 percent wage premium for both undereducated men and women in Portugal.

Duncan and Hoffman (1981) model distinguishes required years of education (E_i^R), years of overeducation (E_i^O) and years of undereducation (E_i^U) such that actual years of education (E_i) is defined:

$$E_i = E_i^R + E_i^O - E_i^U. \quad (3)$$

For overeducated workers, $E_i^O > 0$ and $E_i^U = 0$, while for undereducated workers, $E_i^O = 0$ and $E_i^U > 0$. I apply this decomposition only to statistical and self-declared approach, since the normative approach groups several ISCED educational levels together and, hence, the number of required years is not determined. The components of actual years of education then enter the Duncan and Hoffman's specification as follows:

$$lnY_i = \beta_0 + \beta_1 E_i^R + \beta_2 E_i^O + \beta_3 E_i^U + X_i\gamma + \varepsilon_i. \quad (4)$$

Coefficients β_2 and β_3 (returns to a year of education beyond and below required, respectively) differs from those in Verdugo and Verdugo model as here they have to be interpreted relative to co-workers who have just the required education, in other words, relative to workers in the same occupation but with different (matched) education.

Table 5 Wage models: Mincer and Duncan–Hoffman

	Statistical				Self-declared
	SILC			PIAAC	PIAAC
	Total	*Male*	*Female*	Total	Total
Years of education	0.064*	*0.051**	*0.080**	0.052*	0.052*
R^2	0.40	*0.34*	*0.43*	0.40	0.40
Years of required education	0.092*	*0.072**	*0.118**	0.095*	0.062*
Years of overeducation	0.040*	*0.035**	*0.047**	0.013+	0.025*
Years of undereducation	-0.051*	*-0.035**	*-0.063**	-0.056*	-0.030*
R^2	0.42	*0.35*	*0.47*	0.44	0.43
H_0: $\beta_1 = \beta_2 = -\beta_3$	101.20*	*33.11**	*92.43**	71.36*	73.46*
H_0: $\beta_2 = \beta_3 = 0$	202.11*	*63.44**	*153.69**	47.90*	26.62*

Source: National SILC (Životní podmínky) 2011, PIAAC. Author's computations.
Notes: Weighted. Unweighted sample size: SILC – 6,471 (3,511 men, 2,960 women); PIAAC – 2,034 (and 2,022 in self-declared approach).
* statistically significant at the 1% level. + statistically significant at the 5% level. All regressions include a dummy for sex, experience and experience squared, a dummy for Prague region, a dummy for supervisory position, a dummy for unlimited job contract, a dummy for large company size and nine dummies for industry. Male and female wage equations include also a dummy for parenthood and a dummy for marriage (as they have an opposite effect on wages of men and women).

The Duncan and Hoffman specification has the advantage that two competing theories can be tested here (Hartog and Oosterbeek, 1988). In case that human capital theory holds, wages are not influenced by particular job requirements and, hence, years of required, over-, and under-education should be rewarded equally. Therefore, then $\beta_1 = \beta_2 = -\beta_3$ and equation (4) would be reduced to standard Mincerian wage equation (2). The last but one row of Table 5, which states the results of equations (2) and (4) using OLS regression, shows that based on computed F-statistics the null hypothesis that $\beta_1 = \beta_2 = -\beta_3$ can be rejected notwithstanding the approach and data used.

According to job competition theory (Thurow, 1975), wages are related to a job rather than to a worker and thus only required years of education are rewarded. In such case $\beta_2 = \beta_3 = 0$ and only β_1 in equation (4) would be non-negative. The last row of Table 5 shows that based on computed F-statistics the null hypothesis that $\beta_2 = \beta_3 = 0$ can be also rejected regardless the approach and data used.

The Mincerian-type wage equation shows a 6.4 percent increase of wage per additional year of actual education in SILC and a 5.2 percent increase in PIAAC data. The returns to actual years of education are higher for women than for men. The Duncan and Hoffman model confirms the expectations coming from existing empirics (for instance, Alba-Ramírez, 1993; Nieto, 2014, for Spain) that returns to required years of education are higher than returns to actual years of education. Moreover, returns to years of overeducation are positive but smaller than returns to years of required education ($\beta_1 > \beta_2 > 0$) and returns to years of undereducation are negative but smaller than returns to years of required education ($\beta_1 > -\beta_3 > 0$).

The statistical approach resulting from SILC and PIAAC data differs only in the wage premium for years of overeducation. The self-declared approach in PIAAC results in 3.3 pp smaller returns to years of required education, 1.2 pp higher wage premium from years of overeducation

and about 2.6 pp smaller wage penalty from years of undereducation than under the statistical approach. The most obvious gender difference is that the wage penalty from years of under-education is almost twice as high for women than for men.

CONCLUSION

Three approaches to measurement of over- and undereducation defined as a mismatch between required education in occupations and workers' actual educational attainment have been applied in the literature: Normative approach, based on an exogenous definition determined by job analysts, statistical approach, based on observed educational attainments of workers, and self-declared approach, based on workers' self-evaluation of required education.

This study applied all the approaches on Czech data and, similarly as studies on other countries that applied more than one approach, came to the conclusion that the indices of over- and undereducation vary substantially according to the approach applied: the share of overeducation differ up to 17 pp and the share of undereducation up to 13 pp according the approach applied. Obviously, the choice of an approach to educational mismatch analyses is a dilemma, as all the approaches lead to different results and have their critiques and drawbacks.

Moreover, this study shows that also different data sources lead to different conclusions. I applied national Labour force Survey (LFS), national Statistics on Income and Living Conditions (SILC), and Programme for the International Assessment of Adult Competencies (PIAAC) for measurement of educational mismatch. PIAAC is the only data source that can be used for the self-declared approach. The incidence of both over- and undereducation differs up to 4 pp under the same approach but using different data source. PIAAC yields the highest shares of overeducation compared to LFS and SILC data. Regarding undereducation, LFS provides the lowest shares compared to SILC and PIAAC.

SILC and PIAAC data were further used for analyses of determinants of educational mismatch and wage effects. The impacts of various characteristics on the likelihood of being over- and undereducated vary substantially across both data sources and approaches, resulting in opposite effects in some cases. Even the effect of gender differs across the approaches: Men are less likely to be overeducated rather than matched than women under the normative approach but more likely under the statistical approach. Moreover, the effects are more pronounced in PIAAC than in SILC data.

Wage effects of educational mismatch are in accordance with findings for other countries. First, overeducated workers have lower wages and undereducated workers have higher wages compared to workers with the same education but who work in jobs that match their level of education. However, the estimates still rather vary in their size when different approach or data are used. Compared to SILC data, PIAAC data shows up to 4 pp higher wage penalty for being overeducated and about 2 pp higher wage premium for being undereducated. Both the wage penalty and the wage premium are the highest under the normative approach.

Second, overeducated workers earn more and undereducated workers earn less than their co-workers with exactly the required education. Returns to years of required education are

higher than returns to attained years of education. Moreover, returns to years of surplus education are positive but lower than the returns to required years of education and, vice versa, returns to years of deficit education are negative but this penalty is lower than the returns to required years of education. The effects are qualitatively the same regardless of the data source or approach applied, however, PIAAC data indicates lower returns to years of surplus education compared to SILC data.

The results of all wage models, regardless of data sources and approaches applied, indicate that both the human capital model, which assumes equal returns to required, surplus and deficit years of education, and the job competition model, which assumes no returns to surplus and deficit years of education, can be rejected. The findings suggest that wages are related neither solely to a worker's education nor solely to a job but the reality is somewhere in the middle.

The next steps in the analyses of educational mismatch in the Czech labour market would thus be to include heterogeneity in skills of workers with the same education level (and field of study), as the empirical research in more analysed countries suggests. This might help to overcome another potentially important feature of especially tertiary education in the CR which seems to emerge: the quality and diversity of universities.

Another point would be not only to extend the analysis in time and concentrate on changes of educational mismatch over time but also to focus on the persistence of educational mismatch. While the empirics typically confirms a robust impact of years of work experience on the probability of being overeducated, such a negative effect was proved only in few models tested in this study suggesting that overeducation might be a persistent rather than a temporary phenomenon of mainly young workers.

REFERENCES:

Alba-Ramírez, A. (1993). 'Mismatch in the Spanish labour market. Overeducation?', *Journal of Human Resources*, 28, pp. 259–278.

Bauer, T. (2002). 'Educational mismatch and wages: a panel analysis', *Economics of Education Review*, 21, pp. 221–229.

Bevan, S. and Cowling, M. (2007). 'Job matching in the UK and Europe', Research report 25, The Work Foundation. http://dera.ioe.ac.uk/7125/1/research-report-25-job-matching-in-the-uk-and-europe.pdf

Chevalier, A. (2003). 'Measuring over-education', *Economica*, 70, pp. 509–531.

Duncan, G. J. and Hoffman, S. D. (1981). 'The incidence and wage effects of overeducation', *Economics of Education Review*, 1, pp. 75–86.

Freeman, R. (1976). *The overeducated American*, New York: Academic Press.

Gottschalk, P. and Hansen, M. (2003). 'Is the proportion of college workers in non-college jobs increasing?', *Journal of Labor Economics*, 21, pp. 449–471.

Hartog, J. and Oosterbeek, H. (1988). 'Education, allocation and earnings in the Netherlands: overschooling?', *Economics of Education Review*, 7, pp. 185–194.

Hernández, L. and Serrano, L. (2012). 'Overeducation and its effects on wages: a closer look at the Spanish regions', *Investigaciones Regionales*, 2012, pp. 59–90.

Jeong, B., Kejak, M. and Vinogradov, V. (2008). 'Changing composition of human capital. The Czech Republic, Hungary and Poland', *Economics of Transition*, 16, pp. 247–271.

Karakaya, G., Plasman, R. and Rycx, F. (2007). 'Overeducation on the Belgian labour market: evaluation and analysis of the explanatory factors through two types of approaches', *Compare: A Journal of Comparative and International Education*, 37, pp. 513–532.

Kiker, B. F., Santos, M. C., and De Oliveira, M. M. (1997). 'Overeducation and undereducation: evidence for Portugal', *Economics of Education Review*, 16, pp. 111–125.

McGoldrick, K. M. and Robst, J. (1996). 'Gender differences in overeducation: a test of the theory of differential overqualification', *The American Economic Review*, 86, pp. 280–84.

Mincer, A. (1974). *Schooling, experience and earnings*, New York: NBER Press.

Mysíková, M. (2011). 'EU-SILC and its methodological pitfalls: international comparability and income variables', *Data a výzkum – CSA Info*, 5, pp. 147–170 (in Czech).

Nieto, S. (2014). 'Overeducation, skills and wage penalty: Evidence for Spain using PIAAC data', IREA Working Paper 2014/11, Research Institute of Applied Economics, Barcelona. http://www.ub.edu/irea/working_papers/2014/201411.pdf

OECD (2007). *International migration outlook: SOPEMI 2007 edition*, Paris: OECD.

Rosen, S. (1972). 'Learning and experience in the labor market', *Journal of Human Resources*, 7, pp. 326–342.

Rubb, S. (2003). 'Overeducation: a short or long run phenomenon for individuals?', *Economics of Education Review*, 22, pp. 389–394.

Thurow, L. C. (1975). *Generating inequality*, New York: Basic Books.

Van der Meer, P. and Wielers, R. (1996). 'Educational credentials and trust in the labour market', *Kyklos*, 49, pp. 29–46.

Večerník, J. (2013). 'The changing role of education in the distribution of earnings and household income. The Czech Republic, 1988–2009', *Economics of Transition*, 21, pp. 111–133.

Verdugo, R. and Verdugo, N. (1989). 'The impact of surplus schooling on earnings: some additional findings', *Journal of Human Resources*, 24, pp. 629–643.