Ekonomia. Rynek, gospodarka, społeczeństwo
46(2016), s. 145−167
DOI: 10.17451/eko/46/2016/197
ISSN: 0137-3056
www.ekonomia.wne.uw.edu.pl

Educational Mismatches and Earnings in Poland: Are Graduates Penalised for Being Overeducated?

Leszek Wincenciak*

Abstract

Qualification mismatch is defined as the difference between the level of qualifications held by employees and those required by the type of work they do. Basing on Kiker *et al.* (1997), a measure of overeducation and undereducation is proposed on the basis of the ISCO 08 classification of occupations. The dominant education level is determined for a given occupation's 3-digit group on the basis of the distribution of education levels for employees in that occupation. Each individual having exactly the dominant level of education is considered well-matched. Those with higher levels of education are considered overeducated, those with lower levels − undereducated. An extended Mincer wage regression model with Heckman correction for non-random selection is estimated, using LFS data for Poland for the second quarter of 2013. Significant wage penalties are found in cases of overeducation status, along with positive wage premia for being undereducated, this confirming findings to be noted in the literature of other countries. Applying an approach after Duncan and Hoffman (1981), I find significant positive returns to years of overschooling and negative for underschooling. Young participants on the labour market (graduates) are less penalised for being overeducated, which suggests their overeducation is not necessarily a manifestation of lower ability.

Keywords: educational mismatch, overeducation, undereducation, wage premium, Poland
JEL Codes: J24, J31, J41
DOI: 10.17451/eko/46/2016/197

* Faculty of Economic Sciences, University of Warsaw, e-mail: lwincenciak@wne.uw.edu.pl.

1. Introduction

Qualification mismatch is defined as the difference between the level of qualifications held by employees and those required by the type of work they do. It has been the subject of ever-greater interest in recent years in public debate, as well as in the empirical research of labour economics. Questions are asked regarding the effectiveness of the education system and the process of matching human resources with appropriate tasks.

According to OECD (2011) estimates, in 2005 25.3% of workers on average had qualifications higher than those required for their current job, while 22.2% lacked sufficient qualifications. The educational mismatches were measured on a five-point ISCED scale. The countries of Central and Eastern Europe were among those with underqualification on the largest scale, though these studies placed Poland below average in both categories of mismatch. Measures of mismatches based on employees' self-assessments point to a slightly different picture, with 33.5% of employees on average (30.6% in Poland) saying they have qualifications not necessary for their current jobs, while only 13.3% (16.3% in Poland) recognise that they lack qualifications required for their current work.

The concept of the education-labour market mismatch, though referred to quite often, is rather complex. The problem goes beyond simple skills and requirements. Qualifications (*de jure*) can be understood as a bundle of measurable characteristics such as level of education, specialisation, and formal certificates used to determine skills or ability. After Ortiz and Kucel (2010), a measure of overeducation and undereducation is proposed on the basis of a classification of occupations (ISCO 08), using LFS data to check for realized matches. The 80th percentile of the distribution of educational levels within the given occupation are taken as a cut-off point by which to determine overeducation. Here a slightly different approach is proposed, with the concept of the dominant education being determined (after Kiker *et al.* 1997) for a given occupational group based on the distribution of educational levels for employees in a given occupation. Each individual having exactly the dominant level of education is considered well–matched while those with higher levels of education are considered overeducated, and those with lower levels undereducated. This approach may be justified by the fact that the ISCO 08 classification was supposed to reflect both competences and educational levels required to engage in certain occupations. The Duncan and Hoffman (1981) specification is also tested for.

Constructed measures of over- and undereducation are then used as predictors in Heckman models for hourly wages, to check if they are significant factors in explaining wage differences in Poland. Results are compared for the youngest participants on the labour market (the under-30s), for men and women, to see if there are any interesting differences with respect to socio-demographic characteristics of the labour force. The lower value of the wage penalty for overeducationfor gradu-

ates (under 30) will suggest that their overeducation is not necessarily a sign of lower ability, as is implied by the signalling theory of education (Spence 1973), where less-able individuals have a tendency to overschool in order to mimic their more-able colleagues, with the effect that both outcomes and the allocation of resources are inefficient.

2. Literature findings

Qualification mismatches, caused by information barriers, adjustment costs, low territorial mobility of labour or bad general macroeconomic conditions, are thought to be one of the basic factors that can affect the evolution of the education wage premium. Individuals characterised by a significant scale of mismatch are observed to exhibit the so-called wage penalty over those who are well-matched. This phenomenon can lead to a wider wage dispersion among those who have the same level of education formally.

The theoretical grounds for overeducation to exist on the labour market were laid out by the signalling theory of education (Spence 1973). The basic proposition is that the labour market is characterised by imperfect and asymmetric information. Workers invest in education to send a signal to potential employers about their unobservable ability. The costs of education are presumably higher for low-ability workers. However, if costs of education are not too high, it is possible that low-ability workers will tend to overeducate themselves, trying to be as similar as possible to better candidates (which they are not). On the other hand, high-ability workers will have incentives to extend schooling in the hope of distinguishing themselves from others. Another view is proposed by Thurow's theory of job-competition (1975). There are two ordered queues: one for candidates and the other for jobs. Jobs differ in terms of requirements, productivity and offered wages, while candidates differ in various characteristics like education, experience and skills, which together imply the cost that is necessary for an employer to train them to perform a particular job. The higher the position in the queue, the better the jobs and candidates. To jump to a higher position, workers will have incentives to invest in education or training to get ahead of others and increase their relative chances of getting better jobs. If job requirements are fairly stable over time, this job-competition model explains the existence of overeducation as rational behaviour. A testable hypothesis arising from this approach is that only the years of required schooling affect wage levels.

The means by which the scale of overschooling or underschooling is measured represents a very important issue. To find this measure, researchers need to assess requirements for a particular job. The literature in this field distinguishes three main approaches to this, i.e. workers' self-assessments, information based on job descriptions and information based on worker-job matches as they actually exist.

The first approach − workers' self-assessment of the required level of schooling (or years of schooling) for the job they perform at the given time is subjective. The second approach − job analysis − is based on the use of information contained in occupational classifications. The third approach − based on matches actually in effect − uses the method proposed by Verdugo and Verdugo (1989), as well as the variation thereof after Kiker *et al.* (1997). The required amount of schooling is derived from the mean among actually-observed schooling levels within particular occupations. Verdugo and Verdugo used a 3-digit occupation classification, and this became a standard approach in these circumstances. Workers are classified as overeducated or undereducated if their level of schooling differs by more than one standard deviation from the mean. Kiker *et al.* (1997) in fact proposed the use of the mode as opposed to the mean, classifying workers on the basis of the difference between the said mode and the actual level of education held. Differences in definitions are quite substantial and, as Groot and van den Brink (2000) point out, they are very largely responsible for the observed heterogeneity of results.

Concerns about overeducation were raised by the famous book by Freeman entitled *The Overeducated American*, which was published in 1976. Freeman pointed out that the growing oversupply of college graduates not matched by labour demand would drive returns to education down, with the investment in schooling being questioned in the case of many individuals failing to find suitable jobs requiring the skills they possess or have acquired.

The seminal paper by Duncan and Hoffman (1981) triggered new research in the field, which focused on microdata as opposed to aggregates. They introduced an extension to the Mincer wage-regression model, by proposing the use of variables capturing years of education required for a current job, years of overeducation, and years of undereducation. Inclusion of these variables made it possible to estimate wage returns for required schooling and for over- and undereducation (they used PSID 1976 data for the estimates in their model). The effect of their study was that the return to the surplus education (0.029) was half of that to required schooling (0.063), which they interpret as a potential misallocation of educational resources. The return on each year of underschooling was negative 0.042.

Groot and van den Brink (2000) carry out an excellent literature review and meta-analysis of the literature findings on overeducation. They point out that there is a positive impact of labour-force growth on the incidence of overeducation and a negative effect of unemployment on returns to education.

Hartog (2000) uses the procedure after Duncan and Hoffman (1981) to confirm positive returns to years of overeducation (although smaller in magnitude to the years of required schooling) and negative returns to years of undereducation. Allen and van der Velden (2001) show that there are negative wage returns to skill mismatches (underutilisation of skills in current jobs), albeit with a weak wage penalty relating to the education mismatch. Job satisfaction is also affected very nega-

tively by skill mismatches, though not by educational mismatches. This findings confirmed to some extent by Aracil and van der Velden (2008). Those authors also found a positive impact of being overeducated on wages as they summarise that methodological competences pay off even when they are not required. Bauer (2002) uses panel data for Germany for the period 1984−1998 to study the impact of educational mismatch on wages. Following the model after Verdugo and Verdugo (1989), he found a negative wage premium for the overeducated individuals and a positive one for undereducated ones. He also confirms earlier findings by Duncan and Hoffman (1981) with respect to years of required schooling, years of overeducation and years of undereducation. Using the data from a survey of the University of Granada, Salas-Velasco (2006) found that there are negative wage returns to underutilisation of skills and positive ones to skill deficits. Budria and Moro-Egido (2008) investigate the wage inequality in Spain and find that the wage dispersion within educational groups can to some extent be explained by educational mismatches. They find that incorrect qualification and strong mismatches are associated with wage penalties that range from 13% to 27%.

Korpi and Tahlin (2009) examine the impact of educational mismatch on wages and wage growth in Sweden. The empirical analyses are based on cross-sectional and panel data from standard of living surveys in the period 1974−2000. They find no evidence that wage growth is higher for overeducated workers, and conclude that overeducated employees are penalised early, at the stage of entry on to the labour market, and the returns to education do not recover at later stages of their professional careers. Barone and Ortiz (2010) use REFLEX data to conduct an analysis of wage returns to overeducation in countries across Europe. Their finding is that concerns about overeducation are not justified, as there are no clear signs of wage penalties, at least not in Germany, the Czech Republic, Austria and Italy. In Norway, Finland and The Netherlands, there was more pronounced expansion of the HE system, but these countries are reported to be successful in creating growing demand for skilled employees. Barone and Ortiz point to Spain as a negative example, with overeducation posing a serious risk there, associated with negative wage returns. Negative returns to overeducation are not confirmed by Tsai (2010), who investigates this phenomenon for the American labour market using the Panel Study of Income Dynamics for 1979–2005. Use of a numerical approach in the presence of non-classical measurement error in educational mismatch variables leads to a conclusion that there are no significant wage penalties where overeducation status is concerned. McGuinness and Sloane (2011) studied the overeducation impact on wages for graduates in the UK using REFLEX data. They ound substantial wage penalties for overeducation in the case of both sexes, while penalties for overskilling were confined to men only.

Using POLPAN data, Kiersztyn (2013) shows important long lasting effects of overeducation for two decades in Poland at both the macro and micro levels. Po-

tentially, the penalty associated with overeducation has persistent consequences, even affecting the lifetime earnings of workers in Poland.

3. Methodology

3.1. Measuring over- and undereducation

Overeducation and undereducation dummies

The most difficult task in the research on overeducation has been with measurement. The approach proposed here draws on the idea of Ortiz and Kucel (2010), who used the 80th percentile as a cut-off for determining overeducation in the distribution of education within a given occupation. Using LFS data for Poland for the second quarter of 2013 and for each occupational group at the 3-digit level (ISCO 08), we rather employ a procedure proposed by Kiker *et al.* (1997), and define a mode of education level based on the educational distribution among employees within that occupation (the highest share accounted for by a given education level within a 3-digit occupation). Education is measured on a 6-level scale:

- tertiary MA (EDU = 1);
- tertiary BA (EDU = 2);
- secondary vocational (EDU = 3);
- secondary general (EDU = 4);
- vocational primary (EDU = 5);
- primary (EDU = 6).

$$OVEREDUCATION_j = \begin{cases} 1 \; if \; EDU^{dom} > EDU_j \\ 0 \; if \; EDU^{dom} \leq EDU_j \end{cases}$$

$$UNDEREDUCATION_j = \begin{cases} 1 \; if \; EDU^{dom} < EDU_j \\ 0 \; if \; EDU^{dom} \geq EDU_j \end{cases}$$

To each 3-digit occupation, a number for the dominant education level (EDU^{dom}) is assigned as the representation of the required level of education. Two dummy variables are then created for each individual, as follows

An individual is assumed to be overeducated if having a better education (i.e. represented by a lower number) than the dominant one ($OVEREDUCATION_j = 1$), while an undereducated person is someone having a lower level of education than is typical for the occupation ($UNDEREDUCATION_j = 1$). Under the adopted definition, employees with tertiary MA level of education cannot ever be undereducated. For the same reason, employees with primary education cannot ever be overeducated, which seems reasonable, as lower levels of education are not taken into account.

Table A in the Appendix shows the distribution of educational levels within given 3-digit occupations, and a classification for a modal education level.

Years of overeducation and undereducation

The second approach to the problem of measuring overeducation and undereducation relies on the classic paper of Duncan and Hoffman (1981), breaking down the completed years of schooling for the *j*-th individual into three components: – required years of schooling; – years of education; – years of undereducation. In this regard, the following identity holds (after Leuven and Oosterbeek 2011):

$$Y_j = Y_j^R + max\left(0, Y_j - Y_j^R\right) - max\left(0, Y_j^R - Y_j\right).$$

The problem is now transformed into calculating required years of schooling for a given individual holding a particular job. We use here the same categorisation of realised matches, but instead of defining a dominant education level for each 3-digit occupation, we impute the years of education most typical for a given dominant educational level to be attained. All other calculations are straightforward.

3.2. Empirical model

As the work detailed here investigates the impact of overeducation and undereducation on wages, I use an extended Mincer wage regression model (Mincer, 1974) with Heckman correction (Heckman, 1979) for non-random selection. The model is stated in the following form:

$$\ln w_j = x_j\alpha + \beta_1 O_j + \beta_2 U_j + u_{1j},$$

$$y_0 = z_j\delta + u_{2j},$$

$$u_1 \sim N(0, \sigma),$$

$$u_2 \sim N(0,1),$$

$$corr(u_1, u_2) = \rho.$$

The first equation is the wage equation with the log of hourly net wage in the main job as a dependent variable. The (explanatory) variables on the right hand side in the vector x_j include: sex, age (years), age squared, tenure in current job (months), education (six levels defined as above), firm size (5 levels), sector (public, private), class of settlement unit (6 levels), NACE code (2-digit level) and voivodeship (16 geographical units). Additionally, two dummy variables for overeducation and undereducation are included. The expected sign for β_1 is negative, while that for β_2 is positive. The second equation is the probit selection equation

for labour-market participation. The explanatory variables in vector z_j include: sex, age, age squared, education, marital status, number of children aged 5 or less and disability. Correlation of error terms from the two equations indicates a problem of non-random selection. The model is estimated using a maximum likelihood technique.

The same technique is adopted to estimate the wage equation after Duncan and Hoffman (1981) (hereinafter DH81), by including the three variables Y_j^R – required years of schooling; Y_j^O – years of overeducation; and Y_j^U – years of undereducation in place of two dummy variables for the status of overeducation or undereducation. The wage equation then becomes:

$$\ln w_j = x_j\alpha + \gamma_r Y_j^R + \gamma_o Y_j^O \gamma_u Y_j^U + u_{1j}.$$

As is noted by Leuven and Oosterbeek (2011), a convenient feature of this specification is that it allows for testing of the standard Mincer specification as a special case of the above, which is done by not statistically rejecting the hypothesis of $\gamma_r = \gamma_o = -\gamma_u$. All models are estimated in four specifications for: (1) the entire population, (2) the youngest (under 30), (3) men, and (4) women.

4. Data description and statistics

Data for the analysis come from the Labour Force Survey for Poland for the second quarter of 2013, and consist of 86,126 individuals aged 15 or over, of which 37,098 are employed. Figure 1 shows the kernel density functions for log hourly net wage distributions for matched, (vertically) overeducated and (vertically) undereducated.

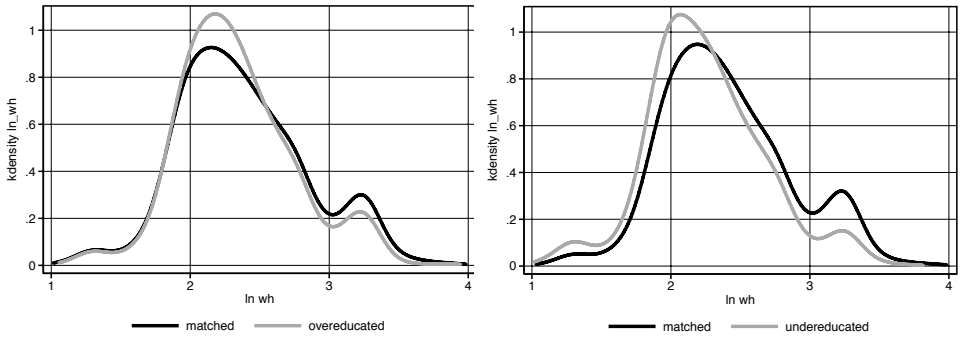

Figure 1. Kernel density of log hourly wage distribution: matched, overeducated and undereducated.

Source: author's own calculations, LFS 2013.

Nearly every fourth employee is considered overeducated (see Table 1) but the same fraction is considered undereducated. Overeducation is more common

among men, the young and those with tertiary education holding a BA degree. In turn, the incidence of undereducation is higher among women, both the young and the old and those with general secondary education. The scale of undereducation among those with primary education is huge, and reflects the fact that only one 3-digit occupation group exhibited dominance in employment of this group of employees.

The education boom which took place in Poland after the transformation to a market economy was reflected in a massive increase in numbers of tertiary graduates within the labour supply. The number of students increased 5-fold between 1991 and 2005 (from under 400,000 to nearly 2 million). At the same time, the share of highly-educated employees in employment also rose, absorbing that growth. It is a statistical issue whether this massive increase resulted in a higher wage penalty for being overeducated, as there was a clearly increased incidence of overeducation in the youngest generation on the labour market. The answer to this question relies on the extent to which overeducation is a manifestation of lower ability, as is implied by the signaling theory of education (Spence 1973).

Descriptive statistics for explanatory variables are presented in Table 2. Reference categories for dummy variables are denoted using asterisks.

Table 1. Incidence of overeducation and undereducation by demographic characteristics of population

Subpopulation	Overeducated (%)	Undereducated (%)
Men	28.25	22.95
Women	21.95	28.68
Age up to 25	35.35	38.39
Age 26−35	33.19	20.51
Age 36−45	23.14	23.22
Age 46−55	19.24	24.72
Age 55+	19.43	33.31
Tertiary MA	18.86	0.00
Tertiary BA	51.85	43.42
Secondary vocational	47.10	16.94
Secondary general	41.99	59.51
Primary vocational	0.11	17.47
Primary	0.00	99.95
Total	24.82	24.93

Source: author's own calculations, LFS 2013.

Table 2. Description of explanatory variables used in the wage model

Variable name	Values	Share (%)
Sex	Man*	54.43
	Woman	45.57
Education	Tertiary, MA degree	23.95
	Tertiary, BA degree	10.42
	Secondary vocational	24.15
	Secondary general*	9.47
	Primary vocational	26.39
	Primary	5.62
Age	Up to 25	8.89
	26−35	29.16
	36−45	27.00
	46−55	22.22
	55+	12.73
Children under 5 in the household	0*	83.14
	1	14.40
	2	2.36
	3 and more	0.10
Disability	Severe	0.21
	Moderate	1.45
	Light	1.38
	None*	96.97
Marital status	Single*	24.90
	Married	68.06
	Widowed	1.91
	Divorced, separated	5.13
Class of settlement unit	Cities 100 th. and more	30.65
	Cities 50−100 th.	9.92
	Cities 20−50 th.	12.39
	Cities 10−20 th.	6.91
	Cities up to 10 th.	5.71
	Rural areas*	34.42
Sector	Public*	32.72
	Private*	67.28
Firm size	Up to 10*	17.85
	Lis.19	18.44
	20−49	16.86
	50−250	25.35
	251 and more	21.50
Current job tenure	Mean	112.34
	Standard dev.	113.07

Source: author's own calculations, LFS 2013.

5. Results

The empirical models described in Section 3.2 were estimated using LFS data for Poland. The most important results are summarised in the table (with the results of OLS estimations presented for comparison in the Appendix). The wage model with overeducation and undereducation dummies reveals a number of interestingresults. First, in Poland, women are less likely to be labour-market participants than men – a finding that is typical for this type of study. Higher education results in a higher probability of labour-market participation, as expected. The presence of a larger number of children aged 5 or under in the household reduces the probability of participation for women, but increases it where men are concerned. Women suffer a wage penalty on the labour market in Poland of magnitude 18.0%. Wages increase with age, assuming peak values at age 51. There are also significant and positive returns to current job tenure, resulting from the accumulation of specific human capital.

Table 3. Estimates of the wage model with over- and undereducation dummies

	Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
Wage equation					
	Sex = woman	–0.1803** [0.000]	–0.1111** [0.000]		
	Age [years]	0.0195** [0.000]	–0.0074 [0.856]	0.0127** [0.000]	0.0361** [0.000]
	Age2	–0.0002** [0.000]	0.0005 [0.540]	–0.0001** [0.002]	–0.0004** [0.000]
	Tenure in current job [months]	0.0006** [0.000]	0.0019** [0.000]	0.0005** [0.000]	0.0006** [0.000]
	EDU=Tertiary MA	0.4206** [0.000]	0.1604** [0.000]	0.3429** [0.000]	0.5390* [0.000]
	EDU=Tertiary BA	0.1652** [0.000]	0.0294 [0.229]	0.1412** [0.000]	0.2182** [0.000]
	EDU=Secondary vocational	0.0326** [0.003]	–0.0196 [0.377]	0.023 [0.164]	0.0507** [0.001]
	EDU=Primary vocational	–0.1709** [0.000]	–0.0904** [0.001]	–0.1737** [0.000]	–0.1881** [0.000]
	EDU=Primary	–0.2970** [0.000]	–0.1308** [0.000]	–0.3071** [0.000]	–0.3225** [0.000]
	Overeducation	**–0.1427**** [0.000]	**–0.0816**** [0.000]	**–0.1348**** [0.000]	**–0.1607**** [0.000]

	Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
	Undereducation	**0.0667****	**0.0161**	**0.0662****	**0.0886****
		[0.000]	[0.445]	[0.000]	[0.000]
	Sector=Private	0.0424**	0.0611**	0.0372**	0.0335*
		[0.000]	[0.019]	[0.005]	[0.014]
	Firm size=11−19	0.0729**	0.0575**	0.0861**	0.0565**
		[0.000]	[0.001]	[0.000]	[0.000]
	Firm size=20−49	0.0939**	0.0979**	0.1103**	0.0710**
		[0.000]	[0.000]	[0.000]	[0.000]
	Firm size=50−250	0.1274**	0.1313**	0.1555**	0.0954**
		[0.000]	[0.000]	[0.000]	[0.000]
	Firm size=251 and more	0.1931**	0.1853**	0.2296**	0.1610**
		[0.000]	[0.000]	[0.000]	[0.000]
	CSU=cities 100 th. and more	0.0641**	0.0929**	0.0781**	0.0501**
		[0.000]	[0.000]	[0.000]	[0.000]
	CSU=cities 50−100 th.	0.011	–0.0168	0.0138	0.0053
		[0.242]	[0.445]	[0.298]	[0.681]
	CSU=cities 20−50 th.	0.0033	0.0329	0.0109	–0.0071
		[0.700]	[0.098]	[0.360]	[0.543]
	CSU=cities 10−20 th.	0.0044	0.0133	0.0238	–0.0151
		[0.669]	[0.576]	[0.097]	[0.280]
	CSU=cities up to 10 th	–0.0226*	0.0243	–0.0248	–0.0218
		[0.040]	[0.357]	[0.108]	[0.155]
	Constant	1.7381**	2.1471**	1.9637**	1.0751**
		[0.000]	[0.000]	[0.000]	[0.000]
Selection equation					
	Sex=woman	–0.5909**	–0.6814**		
		[0.000]	[0.000]		
	Age	0.2200**	0.4696**	0.1645**	0.2741**
		[0.000]	[0.000]	[0.000]	[0.000]
	Age2	–0.0028**	–0.0082**	–0.0023**	–0.0034**
		[0.000]	[0.000]	[0.000]	[0.000]
	EDU=Tertiary MA	0.7344**	0.5780**	0.6033**	0.8195**
		[0.000]	[0.000]	[0.000]	[0.000]
	EDU=Tertiary BA	0.3666**	0.2689**	0.3098**	0.4088**
		[0.000]	[0.000]	[0.000]	[0.000]
	EDU=Secondary vocational	0.2013**	0.1869**	0.2161**	0.2035**
		[0.000]	[0.001]	[0.000]	[0.000]
	EDU=Primary vocational	–0.1158**	–0.1357*	–0.0093	–0.2006**
		[0.000]	[0.016]	[0.835]	[0.000]

	Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
	EDU=Primary	–0.4913**	–0.4881**	–0.3578**	–0.5921**
		[0.000]	[0.000]	[0.000]	[0.000]
	Marital status	0.2661**	0.2529**	0.6682**	–0.1526**
	= married	[0.000]	[0.000]	[0.000]	[0.000]
	Marital status	0.1621**	–0.4970	0.2869**	–0.1530**
	= widowed	[0.000]	[0.505]	[0.001]	[0.003]
	Marital status	0.2090**	–0.2289	0.2054**	0.0048
	= divorced, separated	[0.000]	[0.127]	[0.000]	[0.921]
	Children05=1	–0.3756**	–0.3436**	–0.0078	–0.6216**
		[0.000]	[0.000]	[0.853]	[0.000]
	Children05=2	–0.5287**	–0.6141**	0.3219**	–1.0786**
		[0.000]	[0.000]	[0.002]	[0.000]
	Children05=3	–0.9715**	–0.7834**	0.0017	–7.6200
	and more	[0.000]	[0.022]	[0.996]	[1.000]
	Disability=severe	–1.7462**	–1.6330**	–1.8988**	–1.5236**
		[0.000]	[0.000]	[0.000]	[0.000]
	Disability=moderate	–0.9163**	–0.8258**	–0.9821**	–0.7826**
		[0.000]	[0.000]	[0.000]	[0.000]
	Disability=light	–0.7916**	–0.6565**	–0.9697**	–0.5270**
		[0.000]	[0.000]	[0.000]	[0.000]
	Constant	–3.4975**	–6.1583**	–2.5443**	–4.9279**
		[0.000]	[0.000]	[0.000]	[0.000]
	athrho	–0.2179**	–0.4098**	–0.4329**	0.1365
		[0.000]	[0.001]	[0.000]	[0.194]
	lnsigma	–1.0607**	–1.0688**	–1.0163**	–1.1158**
		[0.000]	[0.000]	[0.000]	[0.000]
	Observations	52813	6196	22852	29961
	Uncensored observations	18249	3266	9680	8569
	Wald chi^2	10712.20	940.99	5330.23	6118.29
	p-value	[0.000]	[0.000]	[0.000]	[0.000]

Note: parameters for NACE section and voivodeships omitted for brevity. Significance levels denoted by: **p-value<0.01; *p-value<0.05.

Source: author's own calculations.

Returns to formal education are substantial and significant. Employees with tertiary education holding MA diplomas earn on average 42.1% more than their colleagues with secondary general education. The returns for youngest employees are significantly lower (only 16.0% over the reference category), which should not be a surprising fact in the light of human capital theory. Returns for MA degrees are also substantially higher for women than for men. Interestingly, there are no significant wage returns for tertiary education with a BA diploma (as set against the reference category) in the case of the youngest employees. The private sector

offers wages that are higher by 3.4 to 6.1% (it offers the greatest premium for graduates). Larger firms also tend to pay more (men exhibit higher returns in relation to firm size than women). The better condition of the labour market in the largest cities is reflected in significantly higher hourly earnings than in rural areas, which is a reference category here.

Overeducated workers exhibit a significantly negative wage premium of 14.3% on average. Performing a job which does not require the level of education possessed is thus associated with a significant wage penalty, as many other studies internationally have also found. The incidence of overeducation is much higher among young generations on the labour market in Poland, but quite surprisingly, the wage penalty is much lower for this group, reaching only 8.2%. A more limited wage penalty among the youngest suggests that their overeducation may not necessarily be a reflection of their consistently lower ability (as would be implied by the signaling theory of education; Spence 1973). This effect may be due to the fact of a normal labour-market history of more frequent job changes and the acquisition of professional experience prior to a move on to jobs to which they are better matched. Furthermore, overeducation seems to penalize women more than men (16.1% and 13.5% respectively).

On the other hand, being undereducated is associated with a significant positive wage premium of 6.7% on average. However, no significant premium for undereducation is found for the youngest workers. Women exhibit a slightly higher wage premium than men, i.e. of 8.9% compared with 6.6%.

Estimates of the DH81 specification are presented in a compact way in Table 4. Parameters for the selection equation and other variables in the wage equation are omitted for brevity. The main focus lies on the estimates of γ_r, γ_o and γ_u.

The results are quite typical for what can be found in the literature on overeducation. The return to each year of required schooling is 11% for the entire population, which seems a little higher than is observed in most developed economies. The return to each year of required schooling is much smaller for graduates (population under 30). Women exhibit higher returns than men, though they still earn less on average. The return to each year of overeducation is roughly 4.6%, which is approximately half that to required years of education (as is commonly found in the literature). However, there are no wage gains for overeducation in the case of graduates. Undereducation brings a return on each year of about 5.6% − slightly less in absolute terms for the youngest and for men.

The hypothesis of the Mincer model being a special case of DH81 specification is strongly rejected in all specifications. Symmetry of returns to years of overeducation and undereducation is not rejected only for men. Thurow's hypothesis of the years of required schooling being the only significant predictors of wages is also rejected by the data, as in all specifications the returns to each year of both underschooling and overschooling are significantly different from zero − excluding the case of the return to overschooling for graduates.

On average, workers who are undereducated (i.e. employed in jobs requiring more schooling than they actually have) earn more than their counterparts who are perfectly matched. This is because the return to each year of undereducation is in absolute terms roughly half of the return to required years of schooling. On the other hand, workers who are overeducated (i.e. employed in jobs requiring fewer years of schooling) earn less on average than their colleagues with the same formal years of schooling, but who do jobs to which they are perfectly matched. This is because the return to each year of required schooling is higher than the return to each year of overeducation (and both are positive). This means that the results of approaches based on both over- and undereducation dummies and ORU specification lead to the same conclusions.

Table 4. Estimates of the wage model with over- and undereducation on DH81 specification

	Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
Wage equation					
	Sex=Woman	–0.1743** [0.000]	–0.1122** [0.000]		
	Age	0.0156** [0.000]	–0.0058 [0.886]	0.0092** [0.001]	0.0327** [0.000]
	Age2	–0.0001** [0.000]	0.0005 [0.559]	–0.0001* [0.042]	–0.0003** [0.000]
	γ_r	0.1104** [0.000]	0.0532** [0.000]	0.1008** [0.000]	0.1313** [0.000]
	γ_o	0.0458** [0.000]	0.0042 [0.507]	0.0419** [0.000]	0.0588** [0.000]
	γ_u	–0.0557** [0.000]	–0.0212** [0.001]	–0.0412** [0.000]	–0.0785** [0.000]
	Observations	52778	6193	22818	29960
	Uncensored observations	18214	3263	9646	8568
	Wald chi^2 p-value	10563.32 [0.000]	950.95 [0.000]	5321.59 [0.000]	5909.41 [0.000]
Tests					
	$\gamma_r = \gamma_o = -\gamma_u$	994.07 [0.000]	88.68 [0.000]	440.56 [0.000]	608.18 [0.000]
	$\gamma_o = -\gamma_u$	6.59 [0.000]	4.02 [0.000]	0.01 [0.000]	13.70 [0.000]

Note: parameters for other variables omitted for brevity.
Significance levels denoted by: **p-value<0.01; *p-value<0.05.
Source: author's own calculations.

6. Conclusions

Qualification mismatches are a sign of sub-optimal allocation of resources in the economy. Inefficiency of this kind may in turn raise questions as to the rationality of spending public money on education. From the point of view of employers, the existence of a large gap between the skills possessed by potential employees and the requirements of the job pose problems of multiple kinds. First, they raise the cost of recruitment. Second, they require some form of testing and screening to disclose the actual productivity of employees. Third, the bridging of gaps requires that major resources be invested in training and acquisition of skills. The considerable costs incurred in adapting employees to employers' expectations, so that theyare able to perform particular jobs, may impair companies' competitiveness, and slow down rates of absorption and implementation of newest technologies.

The problem of mismatches should also be of interest to policymakers. A greater incidence of mismatches may be offer an impulse for reform of the education system to take place, with a view to overall economic efficiency of the process matching human resources being raised through more effective guidance when it comes to the selection of educational paths, and hence more limited waste of the common resources devoted to education.

The methodology adopted for the work detailed in this article was designed to match the specifics of LFS data for Poland, and allowed for the generation of dummy variables for over- and undereducation, on the basis of educational levels and 3-digit occupation groups using a dominant education within each occupation approach. At the same time, the variables for years of required schooling, overschooling and undershooling were created to test the Duncan and Hoffman (1981) specification using data for Poland.

Estimation by way of wage models resulted in significant positive wage returns to undereducation status being found, as well as significant wage penalties associated with being overeducated. The problem of wage penalty for youngest graduates does not seem severe, suggesting that overeducation is not necessarily a sign of lower ability for them, but rather a sign that experience is to be gathered before better-matched jobs are moved on to. Panel data analysis utilising information on job transitions would be desired, for a hypothesis representing a direction to future research to be confirmed.

References

Allen, Jim, and Rolf van der Velden R. 2001. "Educational Mismatches Versus Skill Mismatches: Effects on Wages. Job Satisfaction and On-the-Job Search." *Oxford Economic Papers* 53 (3): 434−452.

Barone, Carlo, and Luis Ortiz. 2011. "Overeducation among European University Graduates: A Comparative Analysis of Its Incidence and the Importance of Higher Education Differentiation. *Higher Education* 61 (3): 325−337.

Bauer, Thomas K. 2002. "Educational Mismatch and Wages: A Panel Analysis." *Economics of Education Review* 21: 221−229.

Bratti, Massimiliano, Robin Naylor, and Jeremy Smith J. 2008. "Heterogeneity in the Returns to Degrees: Evidence from the British Cohort Study 1970." *DEAS, University of Milan, Departamental Working Paper* No. 2008–40.

Budría, Santiago, and Pedro Telhado Pereira. 2005. "Educational Qualifications and Wage Inequality: Evidence for Europe." *IZA Discussion Paper*. No. 1763.

Card, David. 1999. "The Causal Effect of Education on Earning." In: *Handbook of Labor Economics*, Vol. 3, Part A, ed. Orley C. Ashenfelter and David Card, 1801−1863. North Holland: Elsevier.

Chevalier, Arnaud. 2003. "Measuring Over-Education." *Economica* 70: 509−531.

Chevalier, Arnaud, and Joanne Lindley. 2009. "Overeducation and the Skills of UK Graduates." *Journal of the Royal Statistical Society: Series A* 172 (2): 307−337.

Duncan, GGreg J., and Saul D. Hoffman. 1981. "The Incidence and Wage Effects of Overeducation." *Economics of Education Review* 1 (1): 75−86.

Freeman, Richard D. 1976. *The Overeducated American*. New York: Academic Press.

Groot, Wim, and Henriëtte Maasen van den Brink. 2000. "Overeducation in the Labour Market: A Meta-Analysis." *Economics of Education Review* 19 (2): 149−158.

Kiersztyn, Anna. 2013. "Stuck in a Mismatch? The Persistence of Overeducation during Twenty Years of the Post-communist Transition in Poland." *Economics of Education Review* 32: 78−91.

Kiker, B.F, Maria C. Santos, and M. Mendes de Oliveira. 1997. "Overeducation and Undereducation: Evidence for Portugal." *Economics of Education Review* 16 (2): 111−125.

Mavromaras, Kostas G., Seamus McGuinness, Nigel C. O'Leary, Peter J. Sloane, and Zhang Wei. 2010. "Job Mismatches and Labour Market Outcomes: Panel Evidence on Australian University Graduates." *IZA Discussion Paper* No. 5083.

McGuinness, Seamus. 2006. "Overeducation in the Labour Market." *Journal of Economic Surveys* 20: 387−418.

Ortiz, Luis, and Aleksander Kucel. 2008. "Do Fields of Study Matter for Over-Education? The Cases of Spain and Germany." *International Journal of Comparative Sociology* 49 (4−5): 305−327.

Sicherman, Nachum. 1991. “Overeducation in the Labor Market.” *Journal of Labor Economics* 9 (2): 101−122.
Sloane, Peter J., Harminder Battu, and Paul Seaman. 1999. “Overeducation. Undereducation and the British Labour Market.” *Applied Economics* 31 (11): 1437−1453.
Spence, Michael. 1973. “Job Market Signaling.” *Quarterly Journal of Economics*. 87 (3): 355−374.
Thurow, Lester C. 1975. *Generating Inequality: Mechanisms of Distribution in the U.S. Economy*. New York: Basic Books.
Tsai, Yuping. 2010. “Returns to Overeducation: A Longitudinal Analysis of the U.S. Labor Market.” *Economics of Education Review* 29 (4): 606−617.
Verdugo, Richard R., and Naomi Turner Verdugo. 1989. “The Impact of Surplus Schooling on Earnings: Some Additional Findings.” *Journal of Human Resources* 24 (4): 629−643.

Appendix

Table A1. Distribution of education levels within given 3-digit occupations and classification for dominant education level

Occupation code	Tertiary MA (1)	Tertiary BA (2)	Second. vocat. (3)	Second. gen. (4)	Primary vocat. (5)	Primary (6)	Dominant education
111	73.6	16.2	6.3	2.3	0.0	1.6	1
112	65.2	11.6	17.1	3.1	3.0	0.0	1
121	64.6	13.2	14.3	5.3	2.2	0.4	1
122	54.8	17.1	16.8	9.3	2.1	0.0	1
131	19.5	7.8	40.0	12.2	17.0	3.6	3
132	31.3	12.0	35.2	6.7	13.8	1.0	3
133	78.7	17.2	4.1	0.0	0.0	0.0	1
134	76.6	10.6	10.8	1.7	0.4	0.0	1
141	31.5	17.6	34.3	9.2	7.4	0.0	3
142	27.9	18.8	32.1	12.0	8.9	0.3	3
143	37.6	12.0	22.3	11.9	15.3	0.9	1
211	85.7	14.4	0.0	0.0	0.0	0.0	1
212	94.9	5.1	0.0	0.0	0.0	0.0	1
213	89.3	9.1	1.6	0.0	0.0	0.0	1
214	75.6	24.4	0.0	0.0	0.0	0.0	1
215	68.1	31.9	0.0	0.0	0.0	0.0	1
216	63.2	26.1	7.7	1.4	1.6	0.0	1
221	97.2	2.9	0.0	0.0	0.0	0.0	1
222	13.9	35.6	50.2	0.2	0.0	0.1	3
223	17.7	75.6	6.7	0.0	0.0	0.0	2
224	42.7	57.3	0.0	0.0	0.0	0.0	2
225	97.1	2.9	0.0	0.0	0.0	0.0	1
226	100.0	0.0	0.0	0.0	0.0	0.0	1
227	85.9	14.2	0.0	0.0	0.0	0.0	1
228	83.7	14.1	1.7	0.5	0.0	0.0	1
231	96.6	3.4	0.0	0.0	0.0	0.0	1
232	73.5	16.9	9.6	0.0	0.0	0.0	1
233	97.8	2.2	0.0	0.0	0.0	0.0	1
234	93.8	5.7	0.4	0.1	0.0	0.0	1
235	91.3	6.4	1.6	0.8	0.0	0.0	1
241	78.9	16.7	1.6	2.7	0.0	0.0	1
242	80.4	10.7	5.2	3.8	0.0	0.0	1
243	71.8	16.0	7.4	4.9	0.0	0.0	1
244	52.4	22.1	17.1	5.7	2.8	0.0	1
251	76.3	16.9	2.4	4.1	0.0	0.4	1

Occupation code	Tertiary MA (1)	Tertiary BA (2)	Second. vocat. (3)	Second. gen. (4)	Primary vocat. (5)	Primary (6)	Dominant education
252	65.8	28.2	5.4	0.5	0.0	0.0	1
261	99.0	1.0	0.0	0.0	0.0	0.0	1
262	73.1	15.2	9.8	1.9	0.0	0.0	1
263	89.6	5.8	3.5	1.2	0.0	0.0	1
264	84.3	7.2	2.5	5.4	0.5	0.0	1
265	64.2	11.1	13.5	9.8	0.0	1.5	1
311	9.1	11.9	70.0	4.8	4.1	0.0	3
312	12.9	14.8	48.1	6.8	17.4	0.0	3
313	7.1	6.8	50.1	5.4	29.9	0.6	3
314	19.4	10.2	58.3	5.5	4.8	1.9	3
315	47.8	19.9	23.7	3.5	5.2	0.0	1
321	10.2	45.5	41.9	2.5	0.0	0.0	2
322	0.0	21.2	78.8	0.0	0.0	0.0	3
323	100.0	0.0	0.0	0.0	0.0	0.0	1
324	0.0	38.1	61.9	0.0	0.0	0.0	3
325	24.1	34.6	29.8	6.9	4.7	0.0	2
331	43.9	16.6	32.6	6.7	0.2	0.0	1
332	35.7	14.5	31.1	12.8	5.6	0.3	1
333	36.3	11.7	33.6	13.3	4.4	0.8	1
334	44.9	19.1	26.0	10.1	0.0	0.0	1
335	49.3	13.0	26.3	10.2	1.1	0.0	1
341	35.4	14.2	22.2	10.3	17.1	0.9	1
342	59.3	17.5	8.5	6.5	1.7	6.5	1
343	31.6	15.3	25.9	13.7	12.7	0.6	1
351	27.7	13.7	34.3	21.8	2.5	0.0	3
352	24.0	30.9	19.7	18.2	7.3	0.0	2
411	36.1	16.3	31.5	11.8	3.4	1.0	1
412	30.9	26.3	18.3	22.1	2.6	0.0	1
413	23.5	14.6	47.5	14.5	0.0	0.0	3
421	19.9	18.2	36.0	19.3	6.3	0.4	3
422	25.7	19.4	25.1	23.2	6.5	0.2	3
431	27.4	21.9	30.4	15.7	3.7	0.9	3
432	6.0	9.5	40.0	15.7	26.1	2.7	3
441	12.0	14.9	39.8	16.2	16.1	1.0	3
511	42.9	21.2	27.3	0.0	8.6	0.0	1
512	1.6	7.0	33.6	9.9	44.0	3.8	5
513	5.8	11.1	38.2	18.6	22.6	3.7	3
514	3.2	16.0	22.4	16.1	41.1	1.2	5

Occupation code	Tertiary MA (1)	Tertiary BA (2)	Second. vocat. (3)	Second. gen. (4)	Primary vocat. (5)	Primary (6)	Dominant education
515	2.8	6.2	27.9	10.2	42.0	11.0	5
516	20.4	18.4	33.0	10.5	16.3	1.5	3
521	9.5	4.1	26.1	17.3	37.4	5.6	5
522	9.5	10.8	34.2	16.2	26.7	2.6	3
523	6.4	8.4	38.6	24.2	20.8	1.5	3
524	16.7	14.3	30.9	17.2	17.4	3.4	3
531	7.4	13.3	22.2	24.7	25.8	6.5	5
532	4.2	13.3	23.0	11.4	34.6	13.6	5
541	9.4	9.1	33.1	13.0	30.3	5.2	3
611	4.7	5.1	27.0	5.8	42.9	14.5	5
612	2.9	3.4	35.4	5.4	39.7	13.1	5
613	2.1	3.0	22.7	4.3	46.7	21.2	5
621	0.0	0.0	9.6	2.4	46.0	42.0	5
622	27.4	9.6	12.8	0.0	42.6	7.7	5
631	0.9	0.8	12.5	15.1	49.3	21.4	5
633	1.7	2.9	16.0	3.8	46.1	29.5	5
711	1.4	1.1	21.1	4.3	59.9	12.3	5
712	1.2	3.0	25.6	7.7	51.3	11.3	5
713	1.6	1.2	18.4	6.6	62.3	10.0	5
721	0.7	1.8	24.4	7.9	60.7	4.6	5
722	2.5	2.4	28.0	8.1	55.3	3.8	5
723	3.2	3.1	29.5	7.5	53.4	3.3	5
731	1.4	5.3	30.5	11.0	44.9	6.9	5
732	7.3	6.6	43.7	13.1	29.3	0.0	3
741	3.0	5.0	34.9	6.6	49.2	1.4	5
742	9.2	11.8	41.1	9.3	26.6	2.0	3
751	0.8	1.1	14.1	4.1	71.4	8.6	5
752	1.1	4.0	17.0	9.4	56.9	11.6	5
753	0.5	3.60	22.3	7.6	60.8	5.2	5
754	6.1	5.5	27.7	17.1	43.6	0.0	5
811	0.9	2.8	33.3	7.0	49.1	6.9	5
812	0.6	4.5	36.5	5.3	50.0	3.0	5
813	5.5	5.30	40.5	8.4	39.2	1.2	3
814	3.4	7.9	31.7	10.2	42.2	4.6	5
815	1.4	2.8	28.0	6.9	46.3	14.7	5
816	0.0	8.2	26.4	16.7	40.1	8.8	5
817	3.7	4.6	22.3	9.8	48.4	11.2	5
818	1.7	4.6	34.9	8.9	41.7	8.1	5

Occupation code	Tertiary MA (1)	Tertiary BA (2)	Second. vocat. (3)	Second. gen. (4)	Primary vocat. (5)	Primary (6)	Dominant education
821	3.4	4.3	29.0	12.3	43.9	7.2	5
831	3.5	4.8	35.5	7.3	45.9	3.1	5
832	2.8	3.6	37.2	11.7	43.1	4.6	5
833	1.8	3.1	28.3	8.2	53.2	5.5	5
834	0.5	2.1	28.8	5.3	49.5	13.7	5
835	4.9	0.0	35.0	8.7	42.1	9.3	5
911	0.3	2.2	18.1	9.1	49.8	20.6	5
912	2.8	0.0	28.7	15.8	42.4	10.3	5
921	2.6	1.5	9.9	3.8	42.8	39.4	5
931	0.6	1.4	18.3	4.7	50.1	24.9	5
932	3.4	5.5	22.5	9.0	41.1	16.5	5
933	0.7	2.4	24.6	9.5	45.8	17.0	5
941	2.7	2.9	26.6	11.2	47.9	18.9	5
951	0.0	0.0	100.0	0.0	0.0	0.0	3
952	0.0	0.0	0.0	0.0	0.0	100.0	6
961	1.4	0.5	15.5	2.4	47.0	33.3	5
962	0.9	4.9	20.1	7.4	48.3	18.5	5

Table A2. OLS estimation of the wage model with over- and undereducation dummies

Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
Sex=woman	–0.2027**	–0.1594**		
	[0.000]	[0.000]		
Age [years]	0.0303**	0.0411	0.0324**	0.0278**
	[0.000]	[0.285]	[0.000]	[0.000]
Age2	–0.0003**	–0.0004	–0.0004**	–0.0003**
	[0.000]	[0.598]	[0.000]	[0.000]
Tenure in current job [months]	0.0006**	0.0020**	0.0005**	0.0006**
	[0.000]	[0.000]	[0.000]	[0.000]
EDU=Tertiary MA	0.4473**	0.2011**	0.3794**	0.5242**
	[0.000]	[0.000]	[0.000]	[0.000]
EDU=Tertiary BA	0.1818**	0.0552**	0.1633**	0.2111**
	[0.000]	[0.019]	[0.000]	[0.000]
EDU=Secondary vocational	0.0426**	0.0041	0.0399*	0.0467**
	[0.000]	[0.852]	[0.014]	[0.000]

Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
EDU=Primary vocational	–0.1749** [0.000]	–0.0985** [0.000]	–0.1781** [0.000]	–0.1812** [0.000]
EDU=Primary	–0.3210** [0.000]	–0.1736** [0.000]	0.3468** [0.000]	–0.3002 [0.000]
Overeducation	**–0.1429**** [0.000]	**–0.0817**** [0.000]	**–0.1378**** [0.000]	**–0.1599**** [0.000]
Undereducation	**0.0658**** [0.000]	**0.0116** [0.589]	**0.0632**** [0.000]	**0.0897**** [0.000]
Sector=Private	0.0412 [0.000]	0.0520* [0.047]	0.0353** [0.008]	0.0351* [0.010]
Firm size=11−19	0.0727** [0.000]	0.0539** [0.002]	0.0873** [0.000]	0.0550** [0.000]
Firm size=20−49	0.0951** [0.000]	0.0976** [0.000]	0.1139** [0.000]	0.0728** [0.000]
Firm size=50−250	0.1271** [0.000]	0.1299** [0.000]	0.1580** [0.000]	0.0961** [0.000]
Firm size=251 and more	0.1931** [0.000]	0.1868** [0.000]	0.2333** [0.000]	0.1610** [0.000]
CSU=cities 100 th. and more	0.0606** [0.000]	0.0941** [0.000]	0.0697** [0.000]	0.0477** [0.000]
CSU=cities 50−100 th.	0.0095 [0.312]	–0.0163 [0.465]	0.0096 [0.472]	0.0048 [0.710]
CSU=cities 20−50 th.	0.0025 [0.765]	0.0329 [0.101]	0.004 [0.714]	–0.0019 [0.872]
CSU=cities 10−20 th.	0.0025 [0.810]	0.0114 [0.637]	0.0201 [0.165]	–0.0143 [0.310]
CSU=cities up to 10 th.	–0.0236* [0.032]	0.0291 [0.275]	–0.0264 [0.089]	–0.0218 [0.156]
Constant	1.4954** [0.000]	1.4182** [0.004]	1.5285** [0.000]	1.2618** [0.000]
Observations	18408	3299	9765	8643
R2	0.435	0.325	0.397	0.493

Note: parameters for NACE section and voivodeships omitted for brevity. Significance levels denoted by: *p-value<0.01; *p-value<0.05.

Source: author's own calculations.