



UNIVERSITY OF WARSAW

Faculty of Economic Sciences

WORKING PAPERS

No. 28/2016 (219)

LESZEK WINCENCIAK

EDUCATIONAL MISMATCHES AND
EARNINGS: ARE THE GRADUATES MORE
PENALIZED FOR BEING OVEREDUCATED?

WARSAW 2016



UNIVERSITY OF WARSAW

Faculty of Economic Sciences

Educational mismatches and earnings: are the graduates more penalized for being overeducated?

LESZEK WINCENIAK

Faculty of Economic Sciences

University of Warsaw

e-mail: lwincenciak@wne.uw.edu.pl

Abstract

This paper estimates an extended Mincer wage regression model with Heckman correction for non-random selection using LFS data for Poland for the second quarter of 2013. Significant wage penalties are found to overeducation status as well as positive wage premia for being undereducated, which confirms findings that are found in the literature for other countries. Using Duncan and Hoffman (1981) approach, we find significant positive returns to years of overschooling and negative for underschooling. Young participants of the labour market (graduates) are less penalized for being overeducated, which suggests their overeducation is not necessarily a manifestation of lower ability.

Keywords:

educational mismatch, overeducation, undereducation, wage premium, Poland

JEL:

J24, J31, J41

Acknowledgments:

Article prepared as part of the project 'The private value and the rate of return to tertiary education in Poland', funded by the National Science Center, granted by Decision no. 2015/19/B/HS4/03232. I wish to express my gratitude for providing funding by the National Science Center. I wish to thank participants of Warsaw International Conference Meeting 2016 for their valuable comments which helped me to improve this paper. All errors and omissions are entirely mine.

Working Papers contain preliminary research results.

Please consider this when citing the paper.

Please contact the authors to give comments or to obtain revised version.

Any mistakes and the views expressed herein are solely those of the authors.

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 gains more and more interest in recent years in public debate as well as in the empirical research in labour economics. The questions are asked on the effectiveness of the education system and the process of matching of human resources to appropriate tasks.

According to estimates by the OECD (2011) 25.3% of workers on average in 2005 had qualifications higher than those required for the current job, while 22.2% did not have sufficient qualifications. The education mismatches were measured on a five-point ISCED scale. The countries of Central and Eastern Europe were among the ones with the largest scale of underqualification, although Poland by these studies was below average in both categories of mismatch. Measures of mismatches based on self-assessment by employees indicate a slightly different picture — an average of 33.5% of employees (30.6% in Poland) said they have qualifications which are not necessary for the current jobs, while only 13.3% (16, 3% in Poland) have recognized that they lack the qualifications required for current work.

The concept of an education-labour market mismatch, although used 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, specialization, formal certificates that are used to determine the skills or ability. Basing on Ortiz and Kucel (2010) a measure of overeducation and undereducation is proposed based on a classification of occupations (ISCO 08) using LFS data for realized matches. They took the 80th percentile of education levels distribution within given occupation as a cut-off for determining overeducation. Here a slightly different approach is proposed — the concept of dominant education is determined (Kiker et al., 1997) for given occupation group on a basis of the distribution of education levels for employees in given occupation. Each individual having exactly the dominant level of education is considered well matched. Those with higher levels of education are considered overeducated, while those with lower levels are considered undereducated. This approach may be justified by the fact that the ISCO 08 classification was supposed to reflect both competences and education levels required to perform certain occupations. Duncan and Hoffman (1981) specification is also tested.

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 of the labour market (under 30), 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 overeducation for graduates (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 tendency to overschooling to mimic their more able colleagues leading to inefficient outcomes and inefficient allocation of resources.

1 Literature findings

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

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 labour market is characterized 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. If costs of education are not too high it is however possible that the 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 hopes 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 requirements, productivity and offered wages, while candidates differ in various characteristics like education, experience and skills, which imply the cost that is necessary for the employer to train them to perform on a particular job. The higher the position in the queue, the better the jobs and candidates. In order to jump to a higher position, workers will have incentives to invest in education or training to get ahead of the others and to 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. Testable hypothesis resulting from this approach is that only the years of required schooling affect the wage levels.

How to measure the scale of overschooling or underschooling is a very important issue. In order to find this measure researchers need to assess requirements for particular job. The literature of the field distinguishes three main approaches to this problem: workers' self-assessments, information based on job description, information based on realized worker-job matches. The first approach - workers' self-assessment of the required level of schooling (or years of schooling) of the job they presently perform is subjective. The second approach — job analysis — is based on using the information represented in occupational classifications. The third approach — based on realized matches — uses the method proposed by Verdugo and Verdugo (1989) and its variation proposed by Kiker et al. (1997). The required amount of schooling is derived from the mean of actually observed schooling levels within particular occupations. Verdugo and Verdugo used 3-digit occupations classification and it became a standard in this fashion. 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) propose to use a mode instead of the mean and classify workers based on difference between the mode and the actual level of

education held. Differences in definitions are quite substantial and as pointed out by Groot and van den Brink (2000) they are responsible to a large extent for observed heterogeneity of results.

The concerns for overeducation were raised by the famous Freeman's book "The Overeducated American", published in 1976. Freeman pointed out that the growing oversupply of college graduates that is not matched by the labour demand is going to drive the returns to education down, questioning the investment in schooling for many individuals who would not find suitable jobs which would require skills they have.

The seminal paper by Duncan and Hoffman (1981) triggered new research in the field, focusing on a micro data rather than on aggregates. They introduced an extension to Mincer wage regression model by proposing to use variables capturing years of education required for current job, years of overeducation and years of undereducation. Inclusion of these variables made it possible to estimate wage returns to required schooling, to over- and undereducation (they used PSID 1976 data to estimate their model). The effect of their study was that the return to the surplus education (0.029) was half of the return 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) make an excellent literature review and a 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 Duncan and Hoffman (1981) procedure 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), but wage penalty related to education mismatch is weak. Job satisfaction is also strongly negatively affected by skill mismatches but not by education mismatches. This finding is confirmed to some extent by Aracil and van der Velden (2008). In addition to that they find positive impact of being overeducated on wages as they summarize 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 Verdugo and Verdugo (1989) model he finds 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. Salas-Velasco (2006) using the data from a survey of the University of Granada found that there are negative wage returns to underutilization 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 education groups to some extent can be explained by education mismatches. They find that incorrect qualification and strong mismatches are associated with a 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 is based on cross-sectional and panel data from the Level of living surveys 1974-2000. They find no evidence that the wage growth is higher for overeducated workers and conclude that the overeducated employees are penalized early at the entry to the labour market and their returns to education do not recover at the later stages of professional careers. Barone and Ortiz (2010) use REFLEX data to conduct a cross country analysis of wage returns to overeducation in Europe. Their finding is that concerns about overeducation are not justified as there are no clear signs of wage penalties, at least in Germany, Czech Republic, Austria and Italy. In Norway, Finland and Netherlands there was more pronounced expansion of HE system but these countries are reported to be successful in creating growing demand for skilled employees. Barone and Ortiz point out Spain as a negative example, where overeducation is a serious risk, associated with a negative wage returns. Negative returns to overeducation are not confirmed by Tsai (2010) who investigates this phenomenon for the American labour market using Panel Study of Income Dynamics for 1979-2005. Use of numerical approach in the presence of non-classical measurement error in the educational mismatch variables leads to conclusion that there are no significant wage penalties to overeducation status. McGuinness and Sloane (2011) study the overeducation impact on wages for graduates in UK using REFLEX data. They find substantial wage penalties for overeducation for both sexes and for overskilling in the case of men only.

Kiersztyn (2013) using POLPAN data shows important long lasting effects of over-education for two decades in Poland at both, macro and the micro level. The penalty associated with over-education has, potentially, persistent consequences for lifetime earnings of workers in Poland.

2 Methodology

2.1 Measuring the over- and undereducation

Overeducation and undereducation dummies

The most difficult task in the research of overeducation has been the problem of measurement. The approach proposed here draws on idea of Ortiz and Kucel (2010). They were using 80th percentile as a cut-off for determining overeducation in the distribution of education within given occupation. Using LFS data for Poland for the second quarter of 2013 for each occupational group at the 3-digit level (ISCO 08) we rather employ procedure proposed by Kiker et al. (1997) and define a mode of education level based on the education distribution among employees within that occupation (highest share of given education level within 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 ($EDU = 6$).

To each 3-digit occupation a number of a dominant education level (EDU^{dom}) is assigned as the representation of required level of education. Then for every individual the following two dummy variables are created:

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

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

An individual is assumed to be overeducated if having better education (which is represented by lower number) than dominant one ($OVEREDUCATION_j = 1$), while undereducated if having lower education than typical for the occupation ($UNDEREDUCATION_j = 1$). Under 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 A1 in the Appendix shows the distribution of education levels within given 3-digit occupations and a classification for a modal education level.

Years of overeducation and undereducation

Second approach to the problem of measuring overeducation and undereducation relies on classical Duncan and Hoffman paper (1981) and decomposes the completed Y_j years of schooling for j -th individual into three components: Y_j^R — required years of schooling, $Y_j^O = \max(0, Y_j - Y_j^R)$ — years of overeducation, $Y_j^U = \max(0, Y_j^R - Y_j)$ — years of undereducation. The following identity holds (Leuven and Oosterbeek, 2011):

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

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

2.2 Empirical model

In this study we investigate the impact of overeducation and undereducation on wages, therefore an extended Mincer wage regression model (Mincer, 1974) with Heckman correction (Heckman, 1979) for non-random selection is chosen. The model is stated in the following form:

$$\begin{aligned}\ln w_j &= x_j\alpha + \beta_1 O_j + \beta_2 U_j + u_{1j}, \\ y_0 &= z_j\delta + u_{2j}, \\ u_1 &\sim N(0, \sigma), \\ u_2 &\sim N(0, 1), \\ \text{corr}(u_1, u_2) &= \rho.\end{aligned}$$

The first equation is the wage equation with log of hourly net wage in the main job as a dependent variable. The variables on the right hand side (explanatory variables) 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), voivodeship (16 geographical units). Additionally, two dummy variables for overeducation and undereducation are included. Expected sign for β_1 is negative and positive for β_2 . 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 a disability. Correlation of error terms from both equations indicates a problem of non-random selection. The model is estimated using maximum likelihood technique.

The same technique is adopted to estimate Duncan and Hoffman (1981) wage equation (DH81 hereafter) by including three variables: Y_j^R — required years of schooling, Y_j^O — years of overeducation, Y_j^U — years of undereducation in place of two dummy variables for overeducation and undereducation status. 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 Leuven and Oosterbeek (2011) note, a convenient feature of this specification is that it allows to test 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: (1) for entire population, (2) youngest (under 30), (3) men, and (4) women.

3 Data description and statistics

Data for the analysis comes from the LFS for Poland for the second quarter of 2013 and consists of 86 126 individuals aged 15 or more of which 37 098 are employed.

Nearly every fourth employee is considered overeducated (see Table 1). The same amount is considered undereducated. The incidence of overeducation is more common among men, young and those with tertiary education holding BA degree. The incidence of undereducation is more common on the other hand among women, both young and old and those with general secondary education. The scale of undereducation among those with primary education is huge and it reflects the fact that only one 3-digit occupation group exhibited dominance in employment of this group of employees.

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: Own calculations, LFS 2013.

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

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

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
	46-55	22.22
	55+	12.73
Children under 5 in the household	0*	83.14
	1	14.4
	2	2.36
	3 and more	0.1
Disability	Severe	0.21
	Moderate	1.45
	Light	1.38
	None*	96.97
Marital status	Single*	24.9
	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
	11-19	18.44
	20-49	16.86
	50-250	25.35
	251 and more	21.5
Current job tenure	Mean	112.34
	Standard dev.	113.07

Source: Own calculations, LFS 2013.

4 Results

The empirical models described in section 2.2. was estimated using LFS data for Poland. The most important results are summarized in the table below (in the Appendix the results of OLS estimations are presented for comparison). Wage model with overeducation and undereducation dummies reveals a number of interesting

results. First, women are less likely to be labour market participants than men in Poland as is usually found in this type of studies. Higher education results in higher probability of labour market participation, as expected. More kids aged 5 or less in the household reduce the probability of participation for women, but increase it for men.

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

Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
<i>Wage equation</i>				
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]
Undereducation	0.0667** [0.000]	0.0161 [0.445]	0.0662** [0.000]	0.0886** [0.000]
Sector = Private	0.0424** [0.000]	0.0611* [0.019]	0.0372** [0.005]	0.0335* [0.014]
Firm size = 11-19	0.0729** [0.000]	0.0575** [0.001]	0.0861** [0.000]	0.0565** [0.000]
Firm size = 20-49	0.0939** [0.000]	0.0979** [0.000]	0.1103** [0.000]	0.0710** [0.000]
Firm size = 50-250	0.1274** [0.000]	0.1313** [0.000]	0.1555** [0.000]	0.0954** [0.000]
Firm size = 251 and more	0.1931** [0.000]	0.1853** [0.000]	0.2296** [0.000]	0.1610** [0.000]
CSU = cities 100 th. and more	0.0641** [0.000]	0.0929** [0.000]	0.0781** [0.000]	0.0501** [0.000]
CSU = cities 50-100 th.	0.011 [0.242]	−0.0168 [0.445]	0.0138 [0.298]	0.0053 [0.681]
CSU = cities 20-50 th.	0.0033 [0.700]	0.0329 [0.098]	0.0109 [0.360]	−0.0071 [0.543]
CSU = cities 10-20 th.	0.0044 [0.669]	0.0133 [0.576]	0.0238 [0.097]	−0.0151 [0.280]
CSU = cities up to 10 th.	−0.0226* [0.040]	0.0243 [0.357]	−0.0248 [0.108]	−0.0218 [0.155]
Constant	1.7381** [0.000]	2.1471** [0.000]	1.9637** [0.000]	1.0751** [0.000]

Table 3: Cont.

Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
<i>Selection equation</i>				
Sex = woman	−0.5909** [0.000]	−0.6814** [0.000]		
Age	0.2200** [0.000]	0.4696** [0.000]	0.1645** [0.000]	0.2741** [0.000]
Age2	−0.0028** [0.000]	−0.0082** [0.000]	−0.0023** [0.000]	−0.0034** [0.000]
EDU = Tertiary MA	0.7344** [0.000]	0.5780** [0.000]	0.6033** [0.000]	0.8195** [0.000]
EDU = Tertiary BA	0.3666** [0.000]	0.2689** [0.000]	0.3098** [0.000]	0.4088** [0.000]
EDU = Secondary vocational	0.2013** [0.000]	0.1869** [0.001]	0.2161** [0.000]	0.2035** [0.000]
EDU = Primary vocational	−0.1158** [0.000]	−0.1357* [0.016]	−0.0093 [0.835]	−0.2006** [0.000]
EDU = Primary	−0.4913** [0.000]	−0.4881** [0.000]	−0.3578** [0.000]	−0.5921** [0.000]
Marital status = married	0.2661** [0.000]	0.2529** [0.000]	0.6682** [0.000]	−0.1526** [0.000]
Marital status = widowed	0.1621** [0.000]	−0.4970 [0.505]	0.2869** [0.001]	−0.1530** [0.003]
Marital status = divorced, separated	0.2090** [0.000]	−0.2289 [0.127]	0.2054** [0.000]	0.0048 [0.921]
Children05 = 1	−0.3756** [0.000]	−0.3436** [0.000]	−0.0078 [0.853]	−0.6216** [0.000]
Children05 = 2	−0.5287** [0.000]	−0.6141** [0.000]	0.3219** [0.002]	−1.0786** [0.000]
Children05 = 3 and more	−0.9715** [0.000]	−0.7834* [0.022]	0.0017 [0.996]	−7.6200 [1.000]
Disability = severe	−1.7462** [0.000]	−1.6330** [0.000]	−1.8988** [0.000]	−1.5236** [0.000]
Disability = moderate	−0.9163** [0.000]	−0.8258** [0.000]	−0.9821** [0.000]	−0.7826** [0.000]
Disability = light	−0.7916** [0.000]	−0.6565** [0.000]	−0.9697** [0.000]	−0.5270** [0.000]
Constant	−3.4975** [0.000]	−6.1583** [0.000]	−2.5443** [0.000]	−4.9279** [0.000]
$\tanh^{-1} \rho$	−0.2179** [0.000]	−0.4098** [0.001]	−0.4329** [0.000]	0.1365 [0.194]
$\ln \sigma$	−1.0607** [0.000]	−1.0688** [0.000]	−1.0163** [0.000]	−1.1158** [0.000]
Observations	52 813	6 196	22 852	29 961
Uncensored observations	18 249	3 266	9 680	8 569
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 sections and voivodeships omitted for brevity.

Significance levels denoted by: ** p-value < 0.01, * p-value < 0.05.

Women suffer a wage penalty on the labour market in Poland. The magnitude is 18.0%. Wages increase with age, attaining its maximum at the age of 51. There are

also significant and positive returns to current job tenure, resulting from accumulation of specific human capital. 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 to MA degree are also substantially higher for women than for men. Interestingly, there are no significant wage returns to tertiary education with BA diploma (with respect to the reference category) for youngest employees. Private sector offers higher wages by 3.4 to 6.1% (it offers greatest premium for graduates). Larger firms also tend to pay more (men exhibit higher returns to firm size than women). Better labour market condition in the largest cities is reflected in significantly higher hourly earnings than in rural areas, which is a reference category here.

Overeducated workers exhibit 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 significant wage penalty as is found in many other international studies. The incidence of overeducation is much more common among the young generations on the labour market in Poland, but quite surprisingly, the wage penalty is much lower for this group, reaching only 8.2%. Lower wage penalty among the youngest suggests that their overeducation may not necessarily be a reflection of their consistent lower ability (which is implied by signalling theory of education; Spence, 1973). This effect may be due to the fact of normal labour market history of more frequent job changes and gaining professional experience before moving 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 significant positive wage premium of 6.7% on average. No significant premium for undereducation is however found for the youngest workers. Women exhibit slightly higher wage premium than men, that is 8.9% compared to 6.6%.

Estimates of the DH81 specification are presented in a compact way in Table 4. Parameters for selection equation and other variables in 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 of overeducation. Return to each year of required schooling is 11% for entire population, which seems a little higher than what is observed for most of the developed economies. The return to each year of required schooling is much smaller for graduates (population under 30). Women exhibit higher returns than men, although they still do earn less on average. The return to each year of overeducation is roughly 4.6%, which is approximately half the return to required years of education (as is commonly found in the literature). There are however no wage gains for overeducation for the 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 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 only significant predictors of the wage is also rejected by the data, as in all specifications both return to each year of underschooling and to each year of overschooling are significantly different from zero — excluding the case of return to overschooling for graduates.

Table 4: Estimates of the wage model with over- and undereducation in DH81 specification

Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
<i>Wage equation</i>				
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	52 778	6 193	22 818	29 960
Uncensored observations	18 214	3 263	9 646	8 568
Wald chi ²	10563.32	950.95	5321.59	5909.41
p-value	[0.000]	[0.000]	[0.000]	[0.000]
<i>Tests</i>				
$\gamma_r = \gamma_o = -\gamma_u$	994.07 [0.000]	88.68 [0.000]	440.56 [0.000]	608.18 [0.000]
$\gamma_o = -\gamma_u$	6.59 [0.0102]	4.02 [0.0449]	0.01 [0.9043]	13.70 [0.0002]

Note: parameters for other variables omitted for brevity.

Significance levels denoted by: ** p-value < 0.01, * p-value < 0.05.

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

5 Conclusions

Qualification mismatches are a sign of sub-optimal allocation of resources in the economy. The inefficiency of this kind may in turn raise the question of 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. Thirdly, they require large resources to be invested in training and acquisition of skills to bridge the gaps. Significant costs to adapt the employees to employers' expectations to perform particular jobs may result in a deterioration of the competitiveness of companies and may slow down the absorption and implementation of the newest technologies.

The problem of mismatches should also be of interest of policy makers. High scale of mismatches may be an impulse for the reforms of the education system, which aim should be to increase overall economic efficiency of the process of matching human resources through more effective guidance in the selection of educational paths and reduce the waste of common resources devoted to education.

Methodology adopted in this article was designed to the specifics the LFS data for Poland and allowed to generate dummy variables for over- and undereducation based on education 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 underschooling were created to test the Duncan and Hoffman (1981) specification using data for Poland.

Estimation of wage models resulted in finding significant positive wage returns to undereducation status and also significant wage penalties associated with being overeducated. The problem of wage penalty for youngest graduates does not seem to be severe, which suggests to reflect that overeducation for them is not necessarily a sign of lower ability, but rather a sign for gathering experience before moving to better matched jobs. Panel data analysis utilizing information of job transitions would be desired to confirm this hypothesis which is a direction for future research.

Literature

- Allen J., van der Velden R. (2001). Educational Mismatches Versus Skill Mismatches: Effects on Wages, Job Satisfaction and On-the-Job Search, *Oxford Economic Papers*, Vol. 53, No. 3, pp. 434-452.
- Barone, C., Ortiz, L. (2011). Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation. *Higher Education*, 61(3), 325-337.
- Bauer T. (2002). Educational Mismatch and Wages: A Panel Analysis, *Economics of Education Review*, Vol. 21, pp. 221-229.
- Budría S., Pereira P. (2005). Educational Qualifications and Wage Inequality: Evidence for Europe, IZA Discussion Paper, No. 1763, Bonn.
- Chevalier A. (2003). Measuring Over-Education, *Economica*, Vol. 70, pp. 509-531.
- Chevalier A., Lindley J. (2009). Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society: Series A*, 172:307-337.
- Duncan G., Hoffman S. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1):75-86.
- Freeman R. (1976). *The Overeducated American*. Academic Press.
- Groot W., van den Brink H. M. (2000). Overeducation in the Labour Market: A Meta Analysis, *Economics of Education Review*, Vol. 19, pp. 149-158.
- Kiersztyn, A. (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., Santos M., de Oliveira M. (1997). Overeducation and undereducation: evidence for Portugal. *Economics of Education Review*, 16(2):111-125.
- McGuinness S. (2006). Overeducation in the Labour Market, *Journal of Economic Surveys*, Vol. 20, pp. 387-418.
- Ortiz L., Kucel A. (2008). Do fields of study matter for over-education? The cases of Spain and Germany. *International Journal of Comparative Sociology*, 49 (4-5), 305-327.
- Sicherman N. (1991). Overeducation in the Labor Market, *Journal of Labor Economics*, Vol. 9, pp. 101-122.
- Sloane P., Battu H., Seaman P. (1999). Overeducation, Undereducation and the British Labour Market, *Applied Economics*, Vol. 31, pp. 1437-1453.
- Spence M. (1973). Job market signalling. *Quarterly Journal of Economics*, 87: 355-374.
- Thurow L. (1975). *Generating inequality*. Basic Books.
- Tsai Y. (2010). Returns to Overeducation: A Longitudinal Analysis of the US Labor Market, *Economics of Education Review*, Vol. 29, pp. 606-617.
- Verdugo R., Verdugo N. (1989). The Impact of Surplus Schooling on Earnings: Some Additional Findings, *Journal of Human Resources*, Vol. 24, pp. 629-643.

Appendix

Table 1: Distribution of education levels within given 3-digit occupations and a classification for a 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
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

Table 1: Cont.

Occupation Code	Tertiary MA (1)	Tertiary BA (2)	Second. vocat. (3)	Second. gen. (4)	Primary vocat. (5)	Primary (6)	Dominant education
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	18.2	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
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

Table 1: Cont.

Occupation Code	Tertiary MA (1)	Tertiary BA (2)	Second. vocat. (3)	Second. gen. (4)	Primary vocat. (5)	Primary (6)	Dominant education
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	3.6	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	5.3	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
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	34.2	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	18.5	5
933	0.7	2.4	24.6	9.5	45.8	17.0	5
941	2.7	2.9	16.6	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 2: OLS estimates of the wage model with over- and undereducation dummies

Variables	(1) All	(2) Age < 30	(3) Men	(4) Women
Sex = woman	-0.2027** [0.000]	-0.1594** [0.000]		
Age [years]	0.0303** [0.000]	0.0411 [0.285]	0.0324** [0.000]	0.0278** [0.000]
Age ²	-0.0003** [0.000]	-0.0004 [0.598]	-0.0004** [0.000]	-0.0003** [0.000]
Tenure in current job [months]	0.0006** [0.000]	0.0020** [0.000]	0.0005** [0.000]	0.0006** [0.000]
EDU = Tertiary MA	0.4473** [0.000]	0.2011** [0.000]	0.3794** [0.000]	0.5242** [0.000]
EDU = Tertiary BA	0.1818** [0.000]	0.0552* [0.019]	0.1633** [0.000]	0.2111** [0.000]
EDU = Secondary vocational	0.0426** [0.000]	-0.0041 [0.852]	0.0399* [0.014]	0.0467** [0.001]
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]	0.0044 [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	18 408	3 299	9 765	8 643
R ²	0.435	0.325	0.397	0.493

Note: parameters for NACE sections and voivodeships omitted for brevity.

Significance levels denoted by: ** p-value < 0.01, * p-value < 0.05.



FACULTY OF ECONOMIC SCIENCES
UNIVERSITY OF WARSAW
44/50 DŁUGA ST.
00-241 WARSAW
WWW.WNE.UW.EDU.PL