Iceland Liechtenstein Norway grants

Norway grants

instytut badań strukturalnych

IBS WORKING PAPER 03/2020
MAY 2020

OVEREDUCATION AND WAGES: THE ROLE OF COGNITIVE SKILLS AND PERSONALITY TRAITS

Marta Palczyńska

ibs working paper 03/2020
may 2020

OVEREDUCATION AND WAGES: THE ROLE OF COGNITIVE SKILLS AND PERSONALITY TRAITS•

Marta Palczyńska♦

Abstract

This article investigates the role of personality traits and cognitive skills as potential determinants of overeducation and in explaining overeducation wage penalty. Using a representative survey of the Polish working-age population, with well-established measures of cognitive skills and personality traits, I find that accounting for personality and cognitive skills does not change the size and the statistical significance of overeducation wage penalty estimates. My results also demonstrate that personality is one of the contributors to the risk of being overeducated among workers aged 18 to 29 but not among people aged 30 to 68. Among younger workers agreeable individuals are more likely to be overeducated while conscientious ones are less likely. Moreover, lower cognitive skills are associated with the probability of being overeducated.

Keywords: overeducation, educational mismatch, wages, personality traits, cognitive skills, numeracy

JEL: D91, I26, J24, J31

• Research financed from the funds granted for science in 2018-2022 for the implementation of an international co-financed project. This publication was created as part of the "Youth Employment PartnerSHIP" project, which seeks to combat unemployment among young people. It measures the effectiveness of instruments to support young people who aren't working or studying in four countries: Poland, Spain, Italy and Hungary. The project is financed by Iceland, Lichtenstein and Norway through the EEA and Norway Grants Fund for Youth Employment.
* I would like to thank Jan Baran and Iga Magda for their insightful comments. I also thank the participants of the 2016 EALE conference and Warsaw International Economic Meeting 2019 for their discussion and remarks.

♦ Institute for Structural Research (IBS), Warsaw School of Economics. E-mail: marta.palczynska@ibs.org.pl.

ISSN 2451-4373 | IBS ®

1. Introduction

In the last three decades many countries experienced fast improvement in human capital. On the one hand, this investment in human capital is viewed as positive for the economic development but, on the other hand, it raises concerns about the capacity of the economy to make good use of the growing highly skilled workforce. Overeducation arises when qualifications of a worker are higher than qualifications required by his or her job. It might have serious consequences both for individuals and the society. The literature provides consistent evidence that overeducation is associated with wage penalties (Caroleo and Pastore, 2018; Korpi and Tåhlin, 2009; Montt, 2017; Rubb, 2003) and job dissatisfaction (Green and Zhu, 2010; Verhaest and Omey, 2006). From the policymakers' perspective, especially in countries where tertiary education is financed by the state, overeducation can be considered as a waste of resources invested in schooling not put to productive use. The problem of the costs of overeducation and the potential policy responses to it is becoming even more relevant as the incidence of overeducation has increased or remained unchanged in almost all of EU countries over the 2001-2011 (McGuinness et al., 2018). The upward trend in the overeducation in Poland is confirmed by Baran (2018) over the years 2006-2014.

This article investigates the role of personality traits and cognitive skills as potential determinants of overeducation and in explaining overeducation wage penalty. While there is a large literature on the incidence and wage consequences of overeducation, analyses accounting for the heterogeneity of skills among workers with the same level of education are sparse. Recently, impact of non-cognitive skills on labour market outcomes has been increasingly recognised (for reviews, see Almlund, Duckworth, Heckman, & Kautz, 2011; Borghans, Duckworth, Heckman, & Ter Weel, 2008). Non-cognitive skills include a variety of character skills, such as personality traits, preferences and motivations. Previous research has provided evidence that non-cognitive skills predict wages (Heckman and Kautz, 2012; Lindqvist and Vestman, 2011). They also influence the probability of becoming overeducated (Blázquez and Budría, 2012). One of the mechanisms explaining this relationship could be the job search effort. Individuals with stronger belief that their actions influence outcomes (so with greater internal locus of control[1]) send more job applications and have higher reservation wage (Caliendo et al., 2015; McGee, 2015). These results might suggest that workers with greater internal locus of control have wider choice of job offers and that they accept positions with requirements below their qualifications less often.

The omitted variable bias is identified as one of the main challenges of the literature aiming at estimating the overeducation wage penalty (Leuven and Oosterbeek, 2011). Workers with a given level of education may differ with respect to skills, motivation or on-the-job training. The main contribution of this paper is to rule out some of these usually unobserved factors: namely, cognitive and non-cognitive skills. The results show that the differences in cognitive and non-cognitive skills among workers with the same level of education do not explain the overeducation wage penalty but they are related to being overeducated. Cognitive skills are measured by standardised achievement tests (collected in the Programme for International Assessment of Adult Competencies – PIAAC) while non-cognitive skills are measured by the Big Five framework[2] often called "the longitude and latitude

[1] Neuroticism, locus of control, self-esteem, and generalised self-efficacy are strongly correlated; and that they are indicators of a higher order trait: namely, core self-evaluation (Judge et al., 2005, 2002).

[2] The Big Five model defines personality on five dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (the opposite of emotional stability) (McCrae and Costa Jr, 1999).

of non-cognitive skills, by which all more narrowly defined skills may be categorized" (Kautz et al., 2014: 9) and Grit scale. Using propensity score matching approach (PSM) allows to check if results obtained using OLS are not attributable to common support problems.

The remainder of this paper is structured as follows. In Section 2 I summarise the relevant literature; Section 3 describes the data used and presents the methodology employed. Section 4 provides the main results and a sensitivity analysis which corroborates my results and Section 5 concludes.

2. Related literature

2.1. Theoretical explanations of overeducation

There are several competing labour market theories explaining the differences in wages between the workers. Human capital theory (Becker, 1964; Schultz, 1961) assumes that wages are solely determined by the productivity of workers. In its simple version, wages depend on the attained education only, so the returns to years of overeducation should be equal to the returns to years of required education. However, empirical research on the earnings effects of mismatch provides robust evidence that the overeducated individuals earn less than well matched persons with the same level of education. Many authors pointed out that individuals with a given level of education have heterogenous endowments of other productive characteristics such as cognitive or non-cognitive skills and it can be a potential explanation of overeducation wage penalty (Chevalier, 2003; Green and McIntosh, 2007). If the concept of human capital is broadened to include often unobserved characteristics such as cognitive and non-cognitive skills or additional vocational training, the observed overeducation penalties are consistent with the human capital model: they can actually generated by the unobserved differences between workers in other dimensions of human capital.

On the other hand, Thurow's job competition theory (Thurow, 1972) assumes that wages are determined by the characteristics of the job only and worker's education only determines his position in the queue for the best jobs. However, the empirical regularities are not explained by Thurow's job competition theory. Assignment models (see for reviews Sattinger, 1993, 2012) are often viewed as better explaining wage consequences of educational mismatch. They assume that wages are determined by both the characteristics of the job and of the worker and are a solution to an allocation problem of heterogeneous workers to heterogeneous jobs.

The career mobility model (Sicherman and Galor, 1990) suggests that workers might choose a position for which they are overeducated if it is compensated by a higher probability to be promoted thanks to the gained on-the-job training and experience. This theory implies that overeducation is a temporary phenomenon. Also the matching theory (Jovanovic, 1979) predicts that overeducation is temporary. Mismatch arises because of imperfect information about the quality of the match. With tenure, the mismatch between a worker and an employer is detected and the worker improves the match by job search.

2.2. Empirical evidence

The research on wage consequences of overeducation uses extensively a modification of Mincerian earnings equation introduced by Duncan & Hoffman (1981). The so-called ORU (Overeducation / Required education / Undereducation) model decomposes actual years of education into three above categories. The empirical studies have very consistent results across geographical locations and time periods analysed: overeducated individuals earn less than well-matched workers with the same level of education but more than well-matched workers in the same type of job. The meta-analysis by Rubb (2003) estimates that an average rate of return to required years of education is 9.6%, the return to a year of overeducation is 5.2% and a penalty for every year of undereducation is -4.8%. Another approach is to estimate a wage penalty for overeducation compared to workers with the same level of education by means of a dummy variable (Verdugo and Verdugo, 1989). The average overeducation wage penalty across studies using a dummy variable is estimated at 15.3% (McGuinness, 2006). Wincenciak (2016) estimated comparable specifications for Poland and got similar results: a 14.3% overeducation wage penalty in the latter model and 4.6%, 11% and -5.6% in the ORU specification respectively.

There are different approaches to overcoming the problem of unobserved heterogeneity when estimating overeducation wage penalty. First, several studies have used fixed effects (FE) models in order to address this problem. Bauer (2002) shows that the wage penalty becomes smaller when controlling for individual fixed effects. However, it was later shown that, when measurement error is accounted for, panel estimates of overeducation on wages are close to ordinary least squares (OLS) (Dolton and Silles, 2008; Verhaest and Omey, 2012). Moreover, the FE approach was criticized because the effects are only identified from a small fraction of individuals changing their educational match and additionally strict exogeneity assumption may not hold as unobserved characteristics are likely to change with the educational match status (Leuven and Oosterbeek, 2011). Second, instrumental variables are used to account for unobserved heterogeneity (Kleibrink, 2016; Korpi and Tåhlin, 2009). As noted by Leuven and Oosterbeek (2011) these instruments are weak and might not fulfil exclusion restrictions. Last, proxies of skills (Allen and Van der Velden, 2001; Chevalier, 2003; Green and McIntosh, 2007) and direct measures of skills (Kleibrink, 2016; Levels et al., 2014; Sohn, 2010) are included to account for differences in ability between individuals with the same level of education. Non-cognitive skills were analysed in the context of the overeducation wage penalty only by Sohn (2010). In his study the inclusion of locus of control in the OLS wage regression has hardly changed the returns to required, over- and undereducation. However, the analysis included only a single dimension of personality - locus of control.

While most of the analyses use OLS, several studies use non-parametric methods. For example, McGuinness (2008) and McGuinness and Sloane (2011) using samples of graduates from Northern Ireland and UK respectively, find that propensity score matching (PSM) estimates of the overeducation wage penalty are close to OLS. Similarly, Lamo and Messina (2010) used PSM assessing the impact of overeducation on wages in the Estonian working-age population. They also conclude that these estimates are in line with those generated by OLS. Comparable results between OLS and PSM suggest that there is no common support problem. However, none of these studies investigate the role of personality traits as a confounding factor and only McGuinness (2008) uses ability measure in his model.

3. Data and Methods

3.1. Data

This study uses the data from the Polish Follow-up Study to the Programme for International Assessment of Adult Competencies (postPIAAC). The dataset combines information for Poland from the international PIAAC study coordinated by OECD and conducted in 2011/2012 with the country follow-up conducted in 2014/2015. The main goals of postPIAAC were to gather longitudinal information on PIAAC respondents in Poland and to collect additional background information not available in the international study. The sample of postPIAAC are PIAAC respondents who live in Poland during the fieldwork of postPIAAC. The sample size is 5224 completed interviews out of 9366 in PIAAC in 2011/2012. The data collection was carried out from October 2014 to February 2015. This implies that the interval between the interviews for an individual respondent is from 2.5 to 3.5 years.

The unique feature of the PIAAC dataset is that it includes direct measures of cognitive skills. The respondents solved exercises in literacy, numeracy and problem solving in technology-rich environments. The last domain was designed only for people with computer experience so instead of the level of skills, group indicators are included in the analysis: no computer experience or failed a basic test, refused to take the assessment on computer, below level 1, level 1, level 2/level 3. The postPIAAC study contains self-report scales of personality: Big Five Inventory - Short (BFI-S) (Gerlitz and Schupp, 2005; John et al., 1991) and the short Grit scale (Grit-S) (Duckworth & Quinn, 2009). Big Five is the most widely accepted model of personality which assumes that personality traits can be organised into five factors: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (the opposite of emotional stability). Grit is defined as perseverance and a passion for pursuing long-term goals (Duckworth et al., 2007).

The analyses are restricted to people in paid employment (excluding self-employed), working more than 12 hours a week who have answered all the relevant questions. In the main specification undereducated workers are excluded from the sample as the focus of this analysis is on overeducation. In a sensitivity analysis undereducated individuals are treated as matched. The overrepresentation of people 19-26 years old in the PIAAC sample allows for separate analysis for younger (18-29 years old) and older (30-68 years old) workers. Younger workers face different constraints in the labour market and have a higher risk of persistent overeducation (Kiersztyn, 2013).

3.2. Educational mismatch measures

Empirical application of the educational mismatch concept requires a measure of the required level of education for a job. There are three ways of measuring the required level of education in the literature. First, some studies rely on evaluation of each occupation by a professional job analyst (JA). This provides exogenous measure of the required education. However, such evaluations are available only for a limited number of countries and are updated infrequently. Next method uses self-assessment of workers (SA) about the level of education required to do or, alternatively, to get their job. It allows for heterogeneous education requirements within the same occupation. The third measure is derived from realized job matches (RM) (Verdugo and Verdugo, 1989). For each occupation the required education level is defined as the mean of completed years of schooling of all workers in the same occupation. A worker is considered as overqualified if their level of education exceeds the mean in their occupation by at least one standard deviation. Kiker, Santos, & Oliveira (1997) proposed an alternative measure based on realized matches. In their approach the required level of education is the mode of completed schooling among people working in the same occupation.

Although the research on overeducation continues since over three decades, there is no agreement on how to measure required education.[3] Measures used in analyses are chosen mainly based on availability, not their properties. Many studies examine whether the results of earnings equation are sensitive to the measure used. They conclude that the results are independent of the measure of the required education used (Chiswick and Miller, 2010;

[3] For a detailed discussion of the advantages and disadvantages of the different measures see: e.g., Hartog (2000) and Leuven & Oosterbeek (2011).

Cohn and Khan, 1995; Groot and Brink, 2000; Kiker et al., 1997). There are only a few papers which make an attempt to validate these measures. Velden & Smoorenburg (2000) evaluated a SA and a JA measure within one wage equation and concluded that self-assessment method gives more accurate estimates. Verhaest & Omey (2006) compared JA, RM and two types of SA measures using the same method of evaluation. They did not indicate one superior method of measurement. According to Hartog (2000), JA is "conceptually superior" but its measurement is often flawed. Therefore, there is no clear guidelines in the literature regarding the appropriate measure of required education.

This analysis uses worker's self-assessment of educational job requirements. The information was collected using the question: "Still talking about your current job: If applying today, what would be the usual qualifications, if any, that someone would need to GET this type of job?". The possible answers correspond to the answers to the question on the highest completed education. These two questions combined give the indicator of the educational mismatch:

$$OVER_i = \begin{cases} 1 \; if \; EDU_i > REQ_i \\ 0 \; if \; EDU_i \leq REQ_i \end{cases}$$

Where EDU_i is individual i's completed level of education and REQ_i is individual i's required level of education by his job.

3.3. Method

The paper examines the role of cognitive and non-cognitive skills in the overeducation wage penalty comparing the estimates from two specifications: without and with these characteristics. In terms of methodological approach, I use OLS and propensity score matching (PSM) methods.

First, I estimate a wage equation using a modification of Mincerian wage equation with a dummy for overeducation[4]:

$$lnw_i = \theta \mathrm{X}_i + \beta \mathrm{OVER}_i + u_i,$$

where w_i is individual i's gross hourly wage, $OVER_i$ is a dummy for overeducation, X_i is a vector of individual characteristics, and u_i denotes the error term. In the base specification the vector of individual characteristics includes: age, age squared, gender, level of education, field of education, a dummy if still in education, experience, experience squared, a dummy if living with partner, a dummy if having kids 0-6 years old, mother's level of education and number of books at home when respondent was 16. In the full specification individual characteristics include additionally numeracy level, ICT skills level and personality traits. These equations are run for the whole sample and separately by age groups (18-29 years old and 30-68 years old) and education (ISCED 3-4 and ISCED 5-6). In case of the base and full specifications for tertiary graduates (ISCED 5-6), a dummy for a private university is added.

Next, I use matching estimation based on the propensity score (Rosenbaum and Rubin, 1983). The propensity score is the conditional probability of receiving the treatment given a vector of pre-treatment covariates. The identifying assumption of this estimation method is unconfoundedness or conditional independence assumption (CIA) which says that conditional on observable variables that influence selection into treatment, treatment status is as if being

[4] The choice not to include the undereducation dummy in the main specification is mainly driven by the aim to estimate OLS and PSM models on the same sample. I estimate the model with two dummies, for overeducation and for undereducation, as a sensitivity analysis.

randomised. It allows to identify the causal effect of the treatment. The parameter of interest in the analysis is the average treatment effect on the treated (ATT). To empirically investigate if often unobserved variables such as cognitive skills and personality traits affect estimators based on CIA assumption I follow analogical approach to Caliendo et al. (2017) who assessed the impact of often unobserved characteristics on the estimated effectiveness of active labour market programs in Germany. I compare the results when these characteristics are included (full specification) and not (base specification).

I estimate the propensity score using a logit model:

$$\Pr(OVER_i = 1 \mid X_i) = \frac{e^{f(X_i)}}{1 + e^{f(X_i)}}$$

where $OVER_i$ is the treatment indicator: working in a job for which one is overeducated. The control group ($OVER_i = 0$) consists of well-matched individuals. I also check the sensitivity of the results when undereducated individuals are included in the control group (with well-matched individuals).

I consider two matching algorithms with different parameters: nearest neighbour (NN) matching (1 and 4 nearest neighbours) and Epanechnikov kernel matching (bandwidths: 0.02, 0.06 and 0.2). Common support is imposed. Matching quality is assessed by comparing the mean absolute standardised bias (MSB) before and after matching and pseudo R^2 from probit estimation of the propensity score on all the variables on raw and matched samples. Standardised bias (SB) for each covariate X is defined as the difference of sample means among the treated and matched controls as a percentage of the square root of the average of sample variances in both groups:

$$SB(x) = \frac{100(\bar{x}_c - \bar{x}_t)}{\sqrt{(s_{xc}^2 + s_{xt}^2)/2}}$$

where $\bar{x}_c$ is the mean of the control group, $\bar{x}_t$ is the mean of the treatment group, s_{xc}^2 the variance of the control group and s_{xt}^2 the variance of the treatment group (Caliendo and Kopeinig, 2008). I present the MSB overall for all variables, for different specifications, before and after matching.

3.4. Educational mismatch incidence

One third of Polish dependent workers are overeducated, while 14% are undereducated (Table 1). Overeducation is more common among workers below 30 years old (43%) and in this group tertiary graduates are more often overeducated (47%).

Table 1. Incidence of educational mismatch

	Total	18-29	30-68	Age 18-29		Age 30-69	
				ISCED 3-4	ISCED 5-6	ISCED 3-4	ISCED 5-6
Undereducated	14%	11%	15%	15%	2%	21%	1%
Matched	53%	46%	55%	43%	51%	50%	65%
Overeducated	33%	43%	30%	42%	47%	30%	34%
Observations	2601	1666	935	846	742	519	374

Source: Own calculations based on postPIAAC data.

In the analysed sample, after removing observations with missing values in the variables of interest and after limiting the sample to well-matched and overeducated individuals, 40% of the population is overeducated (Table 2). Younger individuals (18-29 years old) are more often overeducated than older ones (30-68 years old): 48% and 37% respectively. The education level hardly differentiates the incidence of overeducation.

Table 2. Incidence of overeducation

	Total	18-29	30-68	Age 18-29		Age 30-69	
				ISCED 3-4	ISCED 5-6	ISCED 3-4	ISCED 5-6
Matched	60%	52%	63%	52%	52%	62%	63%
Overeducated	40%	48%	37%	48%	48%	38%	37%
Observations	1600	1092	508	514	547	262	233

Source: Own calculations based on postPIAAC data.

4. Results

There are three main findings. First, personality plays a significant role for selection into being overeducated among younger workers (aged 18-29) but not among people aged 30-68. Among younger workers agreeable individuals are more likely to be overeducated while conscientious ones are less likely. Second, lower cognitive skills are associated with being overeducated and this result is mainly driven by workers aged 30 to 68 years. Third, accounting for personality and cognitive skills does not change the size and the statistical significance of overeducation wage penalty estimates.

Table 3. Selection into overeducation

	Total	18-29	18-29 (ISCED 3-4)	18-29 (ISCED 5-6)	30-68
Numeracy	**-0.268***	-0.061	-0.194	0.047	-0.296
	(0.106)	(0.104)	(0.165)	(0.155)	(0.150)
ICT (no exp. or failed test)	ref.	ref.	ref.	ref.	ref.
ICT (refused test)	0.115	**1.215****	**1.429***	1.165	-0.044
	(0.406)	(0.443)	(0.635)	(0.691)	(0.463)
ICT (below level 1)	-0.465	0.639	0.481	0.954	-0.617
	(0.434)	(0.388)	(0.535)	(0.867)	(0.530)
ICT (level 1)	-0.528	0.613	0.538	0.703	-0.778
	(0.361)	(0.324)	(0.432)	(0.714)	(0.459)
ICT (level 2 & 3)	0.081	0.585	0.374	0.735	0.200
	(0.376)	(0.308)	(0.403)	(0.649)	(0.480)
B5: Conscientiousness	0.022	**-0.442***	-0.071	**-0.737***	0.252
	(0.235)	(0.208)	(0.368)	(0.286)	(0.365)
B5: Extraversion	-0.116	-0.157	-0.058	-0.138	-0.134
	(0.132)	(0.146)	(0.234)	(0.170)	(0.166)
B5: Agreeableness	0.039	**0.536***	0.125	**0.915****	-0.146
	(0.220)	(0.237)	(0.394)	(0.318)	(0.345)
B5: Openness	-0.107	-0.125	-0.224	-0.180	-0.151
	(0.163)	(0.161)	(0.261)	(0.194)	(0.213)
B5: Neuroticism	0.085	-0.008	0.098	-0.097	0.126
	(0.103)	(0.083)	(0.133)	(0.139)	(0.143)
Grit	-0.046	-0.096	-0.079	-0.082	-0.014
	(0.100)	(0.085)	(0.130)	(0.123)	(0.137)
Observations	1600	1092	514	547	508

*Notes: Logit coefficients. Standard errors are in parentheses. ***/**/* indicate statistical significance at the 0.1%/1%/5%-level. Other covariates: age, age squared, gender, level of education, field of education, a dummy if still in education, experience, experience squared, a dummy if living with partner, a dummy if having kids 0-6 years old, mother's level of education and number of books at home when respondent was 16; in 18-29 (ISCED 5-6) subpopulation additionally a dummy if private university.*

Source: Own calculations based on postPIAAC data.

The effects of personality on selection into overeducation among younger workers are driven by the graduates (ISCED 5-6) (Table 3). The directions of the significant personality-overeducation relationships are in line with the results by Blázquez and Budría (2012). Conscientiousness decreases the risk of being overeducated while agreeableness is associated with higher risk of overeducation. Contrary to their results I do not observe relationships between the risk of overeducation and the other personality traits. Also, the level of ICT skills is not related to being overeducated. Only individuals who refused to take the ICT test are more likely to be overeducated than individuals who never used computer of failed the test in the age group 18-29. This might be rather related to the tendency to avoid competition than the ICT skills level itself.

The overall measures of covariate imbalance suggest that the kernel matching algorithm with the bandwidth of 0.06 provide the best balance of covariates after matching (Appendix Table A3). However, there is problem with finding a good balance between treated and controls in three models: base and full for the total population and full for 18-29 (ISCED 3-4). For these models the mean standardised bias is in the range 5-7% so above the 5% threshold for empirical studies (Caliendo & Kopeinig, 2008). Statistics of the quality of the matching for all the used algorithms are reported in Table A3 and the distributions of propensity scores before and after matching in Figure A1 in the Appendix.

The inclusion of cognitive skills and personality traits brings only marginal or no reduction in the overeducation wage penalty (comparing full specification versus base specification). The reduction is the biggest for the 18-29 years old with secondary education: the overeducation penalty is reduced from 14.3% to 12.7% in OLS model and from 11.7% to 10.7% in PSM model. Surprising is the increase in overeducation wage penalty by 2.4 p.p. among workers over 30 years old after including measures of cognitive and non-cognitive skills. When we analyse individuals with different educational levels separately, the effects are more heterogenous in PSM model than in OLS model: the overeducation wage penalty is almost 11% for individuals with secondary education while it is 18% among overeducated tertiary graduates, while in OLS it is 13% and 17% respectively.

As a sensitivity analysis I estimate models where undereducated individuals are treated as matched and included in the control group (Table 5). The main results described above hold: the inclusion of cognitive skills and personality traits brings only marginal or no reduction in the overeducation wage penalty. When undereducated are included in the control group the estimates of the overeducation wage penalty are slightly higher: at most by 1.5 p.p. It is to be expected as undereducated individuals earn more than matched individuals with the same level of education so including them in the control group increases the average wage in the control group.

Table 4. Estimates of the overeducation wage penalty: OLS and PSM results

	Total		18-29		18-29 (ISCED 3-4)		18-29 (ISCED 5-6)		30-68	
	base	full	base	full	base	full	base	full	base	full
OLS	-0.155***	-0.146***	-0.134***	-0.130***	-0.143**	-0.127*	-0.172***	-0.173***	-0.156***	-0.149***
	(-0.033)	(-0.032)	(-0.030)	(-0.030)	(-0.053)	(-0.053)	(-0.032)	(-0.032)	(-0.042)	(-0.042)
PSM	*-0.162****	*-0.165****	-0.145***	-0.140***	- 0.117***	*-0.107**	-0.189***	-0.182***	-0.135**	-0.159***
	(0.023)	*(0.021)*	(0.028)	(0.030)	(0.035)	*(0.043)*	(0.041)	(0.038)	(0.052)	(0.045)

*Notes: PSM is ATT kernel matching estimator (bw=0.06). In italics ATT with MSB>5%. Standard errors are in parentheses and based on bootstrapping with 50 replications. ***/**/* indicate statistical significance at the 0.1%/1%/5%-level.*

Source: Own calculations based on postPIAAC data.

Table 5. Estimates of the overeducation wage penalty: OLS and PSM results (undereducated in the control group)

	Total		18-29		18-29 (ISCED 3-4)		18-29 (ISCED 5-6)		30-68	
	base	full	base	full	base	full	base	full	base	full
OLS	-0.154***	-0.145***	-0.132***	-0.129***	-0.143**	-0.131*	-0.171***	-0.172***	-0.157***	-0.147***
	(-0.031)	(-0.031)	(-0.030)	(-0.030)	(-0.053)	(-0.053)	(-0.031)	(-0.032)	(-0.039)	(-0.039)
PSM	-0.171***	-0.173***	-0.156***	-0.152***	-0.130***	-0.133***	-0.191***	-0.184***	-0.139**	-0.162***
	(-0.019)	(0.023)	(0.027)	(0.030)	(0.038)	(0.042)	(0.034)	(0.033)	(0.045)	(0.047)

*Notes: PSM is ATT kernel matching estimator (bw=0.06). In italics ATT with MSB>5%. Standard errors are in parentheses and based on bootstrapping with 50 replications. ***/**/* indicate statistical significance at the 0.1%/1%/5%-level.*

Source: Own calculations based on postPIAAC data.

5. Discussion and conclusions

This paper aims at investigating if heterogeneity in cognitive and non-cognitive skills explains the likelihood of being overeducated and the overeducation wage penalty. Thus, it extends research on the consequences of overeducation showing that the differences in other dimensions of human capital such as cognitive and non-cognitive skills do not explain the overeducation wage penalty. These characteristics are often unobserved and suspected to be the source of omitted variable bias. This result suggests that it is rather matching inefficiencies which account for the overeducation.

Some personality traits, namely agreeableness and conscientiousness, are related to the probability of being overeducated among workers aged 18-29 but none are related to being overeducated among workers aged 30-68. Lower cognitive skills are associated with being overeducated and this result is mainly driven by workers aged 30 to 68 years. Even though these skills are related to being overeducated, their overall impact of including them in the model estimating overeducation wage penalty is small. This might be related to the fact that only some analysed skills are related to being overeducated and only in some subpopulations, so the overall effects are limited. It can be also the case that these skills are related to overeducation in occupations/industries where wages are rigid and not related to cognitive and non-cognitive skills.

The analysis also shows heterogeneity of the overeducation wage penalty among educational groups: overeducated workers with tertiary education experience higher penalty than overeducated workers with secondary education. Similar results were obtained for skills match in Poland by Liwiński & Pastore (2019) who showed that quality of skills-job match is related to wages of tertiary graduates but not to wages of graduates from secondary education. This is potentially related to the differences in education premium between these levels of education. The differences in wages are bigger between workers with tertiary and secondary education than between workers with secondary and lower than secondary education (Strawiński et al., 2018).

The results also indicate that the differences between OLS and PSM estimates are quite substantial for some subpopulations and even in the opposite directions: for workers with secondary education the PSM estimate of the overeducation wage penalty is smaller than OLS while for workers with tertiary education the PSM estimate is higher than OLS.

This study is not without limitations. The sample sizes are small and do not allow for a more detailed heterogeneity analysis with respect to age. Moreover, the cognitive and non-cognitive skills are not measured pre-treatment (before becoming overeducated). This does not allow me to rule out reverse causality problem in the selection model. There is limited evidence that mismatch causes cognitive decline (de Grip et al., 2008) which could explain the lower cognitive skills of overeducated older workers.

Nevertheless, the results suggest that the heterogeneity in cognitive and non-cognitive skills analysed explain the differences in wages between overeducated and matched only to a very limited extend. Leuven and Oosterbeek (2011) identified the omitted variable bias as one of the main challenges of the overeducation literature. Omitted variables are important as long as we can name potential candidates. Workers with a given level of education can have different skills and abilities, additional training and different work trajectories (Meroni and Vera-Toscano, 2017). This paper rules out some of them: cognitive skills and personality traits. However, the analysis does not include all the potential sources of workers' heterogeneity. Motivation or vocational skills may also be relevant for overeducation consequences. Further research on this issue is needed.

References

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 53, 434–452.

Almlund, M., Duckworth, A.L., Heckman, J., Kautz, T., 2011. Personality Psychology and Economics, in 'Handbook of Economics of Education', Hanushek EA and Machin SJ and Woessmann L edn, Vol. 4. Elsevier BV.

Baran, J., 2018. A side effect of a university boom: rising incidence of overeducation among tertiary educated workers in Poland. Economics and Business Review 4, 41–63.

Bauer, T.K., 2002. Educational mismatch and wages: a panel analysis. Economics of Education Review 21, 221–229.

Becker, G.S., 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press.

Blázquez, M., Budría, S., 2012. Overeducation dynamics and personality. Education Economics 20, 260–283.

Borghans, L., Duckworth, A.L., Heckman, J.J., Ter Weel, B., 2008. The economics and psychology of personality traits. Journal of Human Resources 43, 972–1059.

Caliendo, M., Cobb-Clark, D.A., Uhlendorff, A., 2015. Locus of control and job search strategies. Review of Economics and Statistics 97, 88–103.

Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. Journal of economic surveys 22, 31–72.

Caliendo, M., Mahlstedt, R., Mitnik, O.A., 2017. Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. Labour Economics 46, 14–25.

Caroleo, F.E., Pastore, F., 2018. Overeducation at a Glance. Determinants and Wage Effects of the Educational Mismatch Based on AlmaLaurea Data. Soc Indic Res 137, 999–1032.

Chevalier, A., 2003. Measuring over-education. Economica 70, 509–531.

Chiswick, B.R., Miller, P.W., 2010. Does the choice of reference levels of education matter in the ORU earnings equation? Economics of Education Review 29, 1076–1085.

Cohn, E., Khan, S.P., 1995. The wage effects of overschooling revisited. Labour Economics 2, 67 – 76.

de Grip, A., Bosma, H., Willems, D., van Boxtel, M., 2008. Job-worker mismatch and cognitive decline. Oxf Econ Pap 60, 237–253.

Dolton, P.J., Silles, M.A., 2008. The effects of over-education on earnings in the graduate labour market. Economics of Education Review 27, 125 – 139.

Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: perseverance and passion for long-term goals. Journal of personality and social psychology 92, 1087.

Duckworth, A.L., Quinn, P.D., 2009. Development and validation of the Short Grit Scale (GRIT–S). Journal of Personality Assessment 91, 166–174.

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

Gerlitz, J.-Y., Schupp, J., 2005. Zur Erhebung der Big-Five-basierten persoenlichkeitsmerkmale im SOEP. DIW Research Notes 4, 2005.

Green, F., McIntosh, S., 2007. Is there a genuine under-utilization of skills amongst the over-qualified? Applied Economics 39, 427–439.

Green, F., Zhu, Y., 2010. Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. Oxford Economic Papers 62, 740–763.

Groot, W., Brink, H.M. van den, 2000. Overeducation in the labor market: a meta-analysis. Economics of Education Review 19, 149–158.

Hartog, J., 2000. Over-education and earnings: where are we, where should we go? Economics of education review 19, 131–147.

Heckman, J.J., Kautz, T., 2012. Hard evidence on soft skills. Labour economics 19, 451–464.

John, O.P., Donahue, E.M., Kentle, R.L., 1991. The big five inventory—versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

Jovanovic, B., 1979. Job matching and the theory of turnover. Journal of political economy 87, 972–990.

Judge, T.A., Bono, J.E., Erez, A., Locke, E.A., 2005. Core self-evaluations and job and life satisfaction: the role of self-concordance and goal attainment. Journal of Applied Psychology 90, 257.

Judge, T.A., Erez, A., Bono, J.E., Thoresen, C.J., 2002. Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? Journal of Personality and Social Psychology 83, 693.

Kautz, T., Heckman, J.J., Diris, R., ter Weel, B., Borghans, L., 2014. Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success (Working Paper No. 20749). National Bureau of Economic Research.

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.F., Santos, M.C., Oliveira, M.M. de, 1997. Overeducation and undereducation: Evidence for Portugal. Economics of Education Review 16, 111 – 125.

Kleibrink, J., 2016. Inept or Badly Matched? — Effects of Educational Mismatch in the Labor Market. LABOUR 30, 88–108.

Korpi, T., Tåhlin, M., 2009. Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. Labour Economics 16, 183 – 193.

Lamo, A., Messina, J., 2010. Formal education, mismatch and wages after transition: assessing the impact of unobserved heterogeneity using matching estimators. Economics of Education Review 29, 1086–1099.

Leuven, E., Oosterbeek, H., 2011. Overeducation and mismatch in the labor market. In: Handbook of the Economics of Education. Elsevier, New York, pp. 283–326.

Levels, M., van der Velden, R., Allen, J., 2014. Educational mismatches and skills: new empirical tests of old hypotheses. Oxford Economic Papers 66, 959–982.

Lindqvist, E., Vestman, R., 2011. The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. American Economic Journal: Applied Economics 101–128.

Liwiński, J., Pastore, F., 2019. Are School-Provided Skills Useful at Work? Results of the Wiles Test. Res High Educ.

McCrae, R.R., Costa Jr, P.T., 1999. A five-factor theory of personality. Handbook of personality: Theory and research 2, 139–153.

McGee, A.D., 2015. How the perception of control influences unemployed job search. Industrial & Labor Relations Review 0019793914556245.

McGuinness, S., 2006. Overeducation in the Labour Market. Journal of Economic Surveys 20, 387–418.

McGuinness, S., 2008. How biased are the estimated wage impacts of overeducation? A propensity score matching approach. Applied Economics Letters 15, 145–149.

McGuinness, S., Bergin, A., Whelan, A., 2018. Overeducation in Europe: trends, convergence, and drivers. Oxf Econ Pap 70, 994–1015.

McGuinness, S., Sloane, P.J., 2011. Labour market mismatch among UK graduates: An analysis using REFLEX data. Economics of Education Review 30, 130–145.

Meroni, E.C., Vera-Toscano, E., 2017. The persistence of overeducation among recent graduates. Labour Economics 48, 120–143.

Montt, G., 2017. Field-of-study mismatch and overqualification: labour market correlates and their wage penalty. IZA Journal of Labor Economics 6, 2.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.

Rubb, S., 2003. Overeducation in the labor market: A comment and re-analysis of a meta-analysis. Economics of Education Review 22, 621–629.

Sattinger, M., 1993. Assignment Models of the Distribution of Earnings. Journal of Economic Literature 31, 831–880.

Sattinger, M., 2012. Assignment Models and Quantitative Mismatches. In: Prepared for the Expert Workshop" Skill Mismatch and Firm Dynamics: Integrating Skills with the World of Work," April.

Schultz, T.W., 1961. Investment in Human Capital. The American Economic Review 51, 1–17.

Sicherman, N., Galor, O., 1990. A Theory of Career Mobility. Journal of Political Economy 98, 169–192.

Sohn, K., 2010. The role of cognitive and noncognitive skills in overeducation. Journal of Labor Research 31, 124–145.

Strawiński, P., Majchrowska, A., Broniatowska, P., 2018. Wage Returns to Different Education Levels. Evidence from Poland. Ekonomista 25–49.

Thurow, L.C., 1972. Education and Economic Equality. Public Interest.

Velden, van der R., Smoorenburg, van M., 2000. The measurement of overeducation and undereducation: self-report vs. job-analyst method. Les Cahiers du Lasmas, Serie Seminaire 9–17.

Verdugo, R.R., Verdugo, N.T., 1989. The impact of surplus schooling on earnings: Some additional findings. Journal of Human Resources 629–643.

Verhaest, D., Omey, E., 2006. The impact of overeducation and its measurement. Social Indicators Research 77, 419–448.

Verhaest, D., Omey, E., 2012. Overeducation, undereducation and earnings: further evidence on the importance of ability and measurement error bias. Journal of Labor Research 33, 76–90.

Wincenciak, L., 2016. Educational mismatches and earnings in Poland: are graduates penalised for being overeducated? Ekonomia. Rynek, Gospodarka, Społeczeństwo 145–167.

Appendix

Table A1. Differences in mean characteristics between overeducated and matched individuals, full results

	Mean (overeducated) – Mean (matched)				
	Total	18-29	18-29 (ISCED3-4)	18-29 (ISCED5-6)	30-68
Hourly wages (in PLN)	-2.751***	-2.025***	-1.673*	-3.275***	-2.617**
Hourly wages (in PLN) log	-0.155***	-0.128***	-0.116*	-0.194***	-0.144**
Age (in years)	-2.169*	0.049	0.308	-0.224	-0.942
Gender	0.038	0.039	0.106*	-0.05	0.045
Level of education					
ISCED 0-2	-0.013	-0.023**	0	0	-0.01
ISCED 3C	-0.164***	-0.141***	-0.291***	0	-0.160***
ISCED 3A_voc or 4	0.147***	0.124***	0.245***	0	0.158***
ISCED 3A	0.026	0.025	0.046	0	0.013
Bechelor	0.022	0.076*	0	0.151**	-0.015
Master	-0.017	-0.061*	0	-0.151**	0.014
Private university (for ISCED 5-6)				0.146***	
Field of education					
Teacher training	-0.036	-0.005	0	-0.013	-0.041
Humanities	0.015	0.031	0.057	0.048	-0.006
Social sciences	0.049	0.046	0.073**	0.004	0.048
Science	0.006	-0.012	0.007	-0.031	0.011
Engineering	-0.001	-0.046	-0.104	0.004	0.028
Agriculture and veterinary	0.006	0.009	0	0.01	0.009
Health and welfare	-0.006	0.009	0.050*	-0.036	-0.017
Services	-0.034	-0.032	-0.083*	0.014	-0.031
In formal education	-0.01	-0.066*	-0.088	-0.058	-0.011
Work experience (in years)	-2.737**	0.257	0.213	0.411	-2.1
Living with partner	-0.048	-0.04	0	-0.090*	-0.009
Having children aged 0-6	0.013	-0.031	-0.034	-0.02	0.034
Mother's education					
ISCED 0-2	-0.049	0.022	0.042	0.014	-0.047
ISCED 3C	-0.005	0.024	0.062	0.015	-0.026
ISCED 3A-4	0.057	0.014	-0.06	0.063	0.066
ISCED 5-6	-0.004	-0.060*	-0.045	-0.092*	0.007

Books at home when 16					
<10	0.031	-0.001	0.006	0.005	0.044
11-100	-0.015	-0.022	-0.039	0.008	-0.012
>100	-0.016	0.023	0.033	-0.013	-0.033
Regional unemployement	0.302	1.212**	1.287	1.497**	-0.142
Cognitive skills					
Literacy	0.014	0.069	0.064	-0.027	-0.027
Numeracy	-0.061	-0.014	-0.062	-0.053	-0.091
ICT (no exp. or failed test)	-0.005	-0.040*	-0.044	-0.015	0.018
ICT (refused test)	0.013	0.039	0.044	0.045*	0.018
ICT (below level 1)	-0.027	-0.001	0.001	0.004	-0.034
ICT (level 1)	-0.043	0.022	0.016	0.014	-0.076*
ICT (level 2 & 3)	0.062	-0.02	-0.018	-0.048	0.075
Personality traits					
B5:Conscientiousness	-0.079	-0.09	-0.013	-0.125*	-0.035
B5: Extraversion	-0.158*	-0.088	-0.042	-0.097	-0.189*
B5: Agreeableness	-0.098	-0.023	0.02	-0.034	-0.094
B5: Openness	-0.168*	-0.084	-0.077	-0.057	-0.183
B5: Neuroticism	0.054	-0.035	0.034	-0.107	0.067
Grit	-0.062	-0.03	0.073	-0.142	-0.056
Observations	1600	1092	514	547	508

Source: Own calculations based on postPIAAC data.

Table A2. Selection into overeducation (logit coefficients), full results

	Total	18-29	18-29 (ISCED3-4)	18-29 (ISCED5-6)	30-68
Age	0.020	-0.525	-1.825*	2.464	0.184
	(0.087)	(0.641)	(0.894)	(1.596)	(0.173)
Age ^2	-0.000	0.009	0.036	-0.052	-0.002
	(0.001)	(0.013)	(0.018)	(0.030)	(0.002)
Female	-0.029	0.035	-0.104	0.084	-0.079
	(0.213)	(0.155)	(0.276)	(0.245)	(0.312)
Level of education					
ISCED 0-2	-1.954**	-2.000**			-2.046
	(0.659)	(0.696)			(1.038)
ISCED 3C	-1.739**	-1.765**	-3.140**		-1.855*
	(0.578)	(0.526)	(0.937)		(0.874)
ISCED 3A_voc or 4	0.134	0.710	-0.381		0.013
	(0.478)	(0.363)	(0.771)		(0.791)
ISCED 3A	ref.	ref.	ref.		ref.
Bechelor	-0.086	0.833*		0.470	-0.885
	(0.452)	(0.348)		(0.273)	(0.938)
Master	-0.235	0.245		ref.	-0.288
	(0.456)	(0.366)			(0.758)
In formal education	-0.509	-0.780***	-0.749	-0.674*	-0.362
	(0.264)	(0.220)	(0.395)	(0.285)	(0.738)
Field of education					
Teacher training	-0.320	-0.454		-0.596	-0.152
	(0.366)	(0.411)		(0.482)	(0.501)
Humanities	ref.	ref.	ref.	ref.	ref.
Social sciences	0.268	-0.234	1.632	-0.718	0.471
	(0.386)	(0.329)	(1.024)	(0.410)	(0.563)
Science, mathematics and computer science	0.076	-0.581	0.637	-0.945	0.460
	(0.424)	(0.385)	(0.834)	(0.475)	(0.621)
Engineering	0.295	-0.359	0.510	-0.334	0.582
	(0.389)	(0.345)	(0.764)	(0.448)	(0.575)
Agriculture and veterinary	0.204	-0.117	0.369	0.325	0.441
	(0.640)	(0.476)	(0.981)	(0.644)	(0.860)
Health and welfare	-0.338	-0.428	1.098	-0.975	-0.387
	(0.390)	(0.373)	(0.985)	(0.504)	(0.898)

Services	0.178	-0.509	0.495	-0.351	0.454
	(0.444)	(0.440)	(0.887)	(0.649)	(0.628)
Work experience (in years)	-0.068	0.146**	0.148	0.279**	-0.129*
	(0.045)	(0.052)	(0.088)	(0.100)	(0.063)
Work experience ^2	0.001	-0.004*	-0.003	-0.012	0.002
	(0.001)	(0.002)	(0.002)	(0.006)	(0.001)
Living with partner=1	-0.087	-0.302	-0.174	-0.411	-0.054
	(0.209)	(0.164)	(0.287)	(0.241)	(0.308)
Having kids aged 0-6=1	0.003	-0.034	-0.344	0.207	0.072
	(0.237)	(0.227)	(0.403)	(0.328)	(0.336)
Mother's education					
ISCED 0-2	ref.	ref.	ref.	ref.	ref.
ISCED 3C	0.078	-0.329	-0.201	-0.189	0.119
	(0.252)	(0.287)	(0.445)	(0.579)	(0.314)
ISCED 3A-4	0.263	-0.482	-0.751	0.092	0.492
	(0.246)	(0.307)	(0.467)	(0.548)	(0.295)
ISCED 5-6	0.040	-0.882*	-1.387	-0.368	0.281
	(0.350)	(0.394)	(0.798)	(0.575)	(0.515)
Books at home when 16					
<10	ref.	ref.	ref.	ref.	ref.
11-100	-0.933*	-0.383	-0.299	-0.637	-1.032*
	(0.358)	(0.422)	(0.503)	(1.296)	(0.483)
>100	-0.978*	-0.021	0.134	-0.259	-1.223*
	(0.410)	(0.462)	(0.592)	(1.250)	(0.575)
Local unemployment	0.010	0.045**	0.052	0.050*	-0.008
	(0.021)	(0.016)	(0.033)	(0.020)	(0.031)
Numeracy	-0.268*	-0.061	-0.194	0.047	-0.296
	(0.106)	(0.104)	(0.165)	(0.155)	(0.150)
ICT (no exp or failed test)	ref.	ref.	ref.	ref.	ref.
ICT (refused test)	0.115	1.215**	1.429*	1.165	-0.044
	(0.406)	(0.443)	(0.635)	(0.691)	(0.463)
ICT (below level 1)	-0.465	0.639	0.481	0.954	-0.617
	(0.434)	(0.388)	(0.535)	(0.867)	(0.530)
ICT (level 1)	-0.528	0.613	0.538	0.703	-0.778
	(0.361)	(0.324)	(0.432)	(0.714)	(0.459)
ICT (levels 2 & 3)	0.081	0.585	0.374	0.735	0.200
	(0.376)	(0.308)	(0.403)	(0.649)	(0.480)
B5: Conscientiousness	0.022	-0.442*	-0.071	-0.737*	0.252
	(0.235)	(0.208)	(0.368)	(0.286)	(0.365)

B5: Extraversion	-0.116	-0.157	-0.058	-0.138	-0.134
	(0.132)	(0.146)	(0.234)	(0.170)	(0.166)
B5: Agreeableness	0.039	0.536*	0.125	0.915**	-0.146
	(0.220)	(0.237)	(0.394)	(0.318)	(0.345)
B5: Openness	-0.107	-0.125	-0.224	-0.180	-0.151
	(0.163)	(0.161)	(0.261)	(0.194)	(0.213)
B5: Neuroticism	0.085	-0.008	0.098	-0.097	0.126
	(0.103)	(0.083)	(0.133)	(0.139)	(0.143)
Grit	-0.046	-0.096	-0.079	-0.082	-0.014
	(0.100)	(0.085)	(0.130)	(0.123)	(0.137)
Private university (for ISCED 5-6)				0.599*	
				(0.279)	
Constant	0.837	6.985	22.441*	-30.274	-2.205
	(1.517)	(8.193)	(10.980)	(20.867)	(3.807)
Observations	1600	1092	514	547	508

Source: Own calculations based on postPIAAC data.

Table A3. Matching quality

Total	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	11.95	9.39	0.08	0.04	10.11	6.29	0.09	0.03
NN_4	11.95	5.73	0.08	0.02	10.11	5.72	0.09	0.03
Kernel_0.02	11.95	5.58	0.08	0.02	10.11	5.41	0.09	0.03
Kernel_0.06	11.95	5.64	0.08	0.02	10.11	5.26	0.09	0.02
Kernel_0.2	11.95	6.86	0.08	0.02	10.11	5.83	0.09	0.03
18-29	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	10.80	6.21	0.09	0.02	8.94	6.06	0.10	0.03
NN_4	10.80	3.86	0.09	0.01	8.94	4.00	0.10	0.02
Kernel_0.02	10.80	2.84	0.09	0.00	8.94	3.84	0.10	0.01
Kernel_0.06	10.80	2.73	0.09	0.00	8.94	3.72	0.10	0.01
Kernel_0.2	10.80	3.45	0.09	0.01	8.94	3.25	0.10	0.01
18-29 (ISCED3-4)	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	13.29	7.83	0.13	0.04	10.03	8.79	0.15	0.07
NN_4	13.29	6.24	0.13	0.02	10.03	6.53	0.15	0.03
Kernel_0.02	13.29	5.52	0.13	0.02	10.03	6.36	0.15	0.03
Kernel_0.06	13.29	4.85	0.13	0.01	10.03	6.41	0.15	0.03
Kernel_0.2	13.29	4.75	0.13	0.01	10.03	5.68	0.15	0.03
18-29 (ISCED5-6)	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	11.62	8.50	0.07	0.04	10.74	6.85	0.09	0.03
NN_4	11.62	5.08	0.07	0.02	10.74	5.79	0.09	0.02
Kernel_0.02	11.62	4.46	0.07	0.01	10.74	6.21	0.09	0.01
Kernel_0.06	11.62	3.51	0.07	0.01	10.74	4.82	0.09	0.01
Kernel_0.2	11.62	4.62	0.07	0.01	10.74	4.58	0.09	0.01
30-68	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	10.49	7.42	0.09	0.02	10.84	6.35	0.12	0.05
NN_4	10.49	4.74	0.09	0.01	10.84	6.31	0.12	0.03
Kernel_0.02	10.49	3.89	0.09	0.01	10.84	5.06	0.12	0.03
Kernel_0.06	10.49	3.75	0.09	0.01	10.84	4.30	0.12	0.02
Kernel_0.2	10.49	4.62	0.09	0.01	10.84	4.48	0.12	0.02

Notes: MB_bef: the mean absolute standardised bias before matching; MB_aft: the mean absolute standardised bias after matching; r2bef: Pseudo R^2 from probit estimation of the propensity score on all the variables on raw samples; r2aft: Pseudo R^2 from probit estimation of the propensity score on all the variables on matched samples.

Source: Own calculations based on postPIAAC data.

Table A4. OLS results

	Total		18-29		18-29 (ISCED3-4)		18-29 (ISCED5-6)		30-68	
	base	full	base	full	base	full	base	full	base	full
Overedu.	-0.155***	-0.146***	-0.134***	-0.130***	-0.143**	-0.127*	-0.172***	-0.173***	-0.156***	-0.149***
	(0.033)	(0.032)	(0.030)	(0.030)	(0.053)	(0.053)	(0.032)	(0.032)	(0.042)	(0.042)
Numeracy		0.033		0.044*		0.031		0.037		0.027
		(0.017)		(0.018)		(0.023)		(0.025)		(0.020)
B5: Con		0.099*		0.025		-0.017		0.082		0.125*
		(0.038)		(0.032)		(0.049)		(0.048)		(0.049)
B5: Extr		0.017		0.028		0.040		0.026		0.029
		(0.021)		(0.022)		(0.035)		(0.027)		(0.026)
B5: Agr		-0.119**		-0.051		-0.023		-0.068		-0.141**
		(0.041)		(0.036)		(0.040)		(0.055)		(0.053)
B5: Opn		0.001		0.005		0.008		-0.001		-0.010
		(0.024)		(0.030)		(0.039)		(0.036)		(0.032)
B5: Neu		-0.045**		-0.039*		-0.027		-0.045*		-0.057**
		(0.014)		(0.018)		(0.024)		(0.021)		(0.019)
Grit		-0.016		-0.004		0.010		-0.018		-0.015
		(0.014)		(0.013)		(0.016)		(0.019)		(0.018)
Obs.	1600	1600	1092	1092	514	514	547	547	508	508
R^2	0.412	0.432	0.213	0.236	0.144	0.178	0.207	0.234	0.473	0.498

*Notes: Standard errors are in parentheses. ***/**/* indicate statistical significance at the 0.1%/1%/5%-level. Other covariates: age, age squared, gender, level of education, field of education, a dummy if still in education, ICT levels, experience, experience squared, a dummy if living with partner, a dummy if having kids 0-6 years old, mother's level of education and number of books at home when respondent was 16; in 18-29 (ISCED 5-6) subpopulation additionally a dummy if private university.*

Source: Own calculations based on postPIAAC data.

Table A5. OLS results: specification with undereducation dummy

	Total		18-29		18-29 (ISCED3-4)		18-29 (ISCED5-6)		30-68	
	base	full	base	full	base	full	base	full	base	full
Overedu.	-0.154***	-0.145***	-0.132***	-0.129***	-0.143**	-0.131*	-0.171***	-0.172***	-0.157***	-0.147***
	(0.031)	(0.031)	(0.030)	(0.030)	(0.053)	(0.053)	(0.031)	(0.032)	(0.039)	(0.039)
Underedu.	0.042	0.045	0.013	0.024	-0.010	0.009	0.161	0.165	0.052	0.056
	(0.036)	(0.035)	(0.049)	(0.048)	(0.058)	(0.055)	(0.142)	(0.163)	(0.046)	(0.045)
Numeracy		0.041*		0.039*		0.022		0.035		0.041*
		(0.017)		(0.017)		(0.023)		(0.026)		(0.020)
B5: Con	0.095**		0.029		0.029		0.083		0.116*	0.095**
	(0.036)		(0.032)		(0.051)		(0.048)		(0.045)	(0.036)
B5: Extr	0.015		0.019		0.022		0.025		0.023	0.015
	(0.019)		(0.019)		(0.026)		(0.028)		(0.023)	(0.019)
B5: Agr	-0.120**		-0.065		-0.076		-0.067		-0.139**	-0.120**
	(0.038)		(0.033)		(0.044)		(0.055)		(0.048)	(0.038)
B5: Opn	0.004		0.016		0.022		-0.003		-0.005	0.004
	(0.022)		(0.026)		(0.031)		(0.035)		(0.027)	(0.022)
B5: Neu	-0.042**		-0.032		-0.021		-0.044*		-0.054**	-0.042**
	(0.014)		(0.016)		(0.020)		(0.021)		(0.017)	(0.014)
Grit	-0.011		0.003		0.009		-0.019		-0.012	-0.011
	(0.013)		(0.012)		(0.014)		(0.019)		(0.017)	(0.013)
Obs.	1819	1819	1219	1219	610	610	555	555	600	600
R^2	0.394	0.416	0.202	0.222	0.135	0.162	0.211	0.236	0.449	0.476

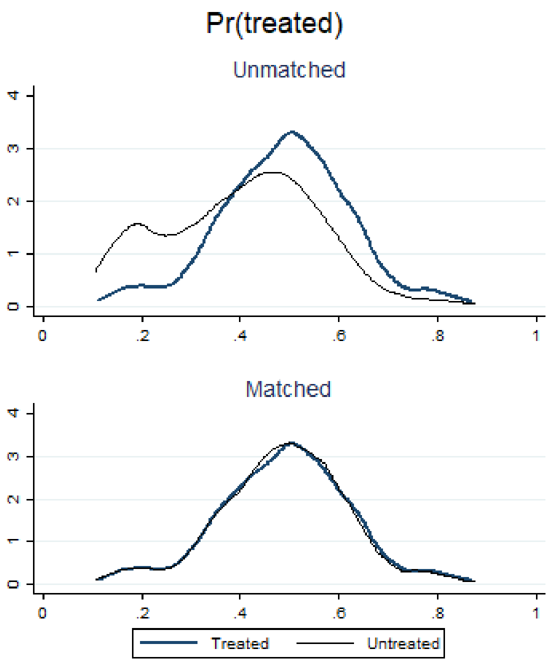
*Notes: Standard errors are in parentheses. ***/**/* indicate statistical significance at the 0.1%/1%/5%-level. Other covariates: age, age squared, gender, level of education, field of education, a dummy if still in education, ICT levels, experience, experience squared, a dummy if living with partner, a dummy if having kids 0-6 years old, mother's level of education and number of books at home when respondent was 16; in 18-29 (ISCED 5-6) subpopulation additionally a dummy if private university.*

Source: Own calculations based on postPIAAC data.

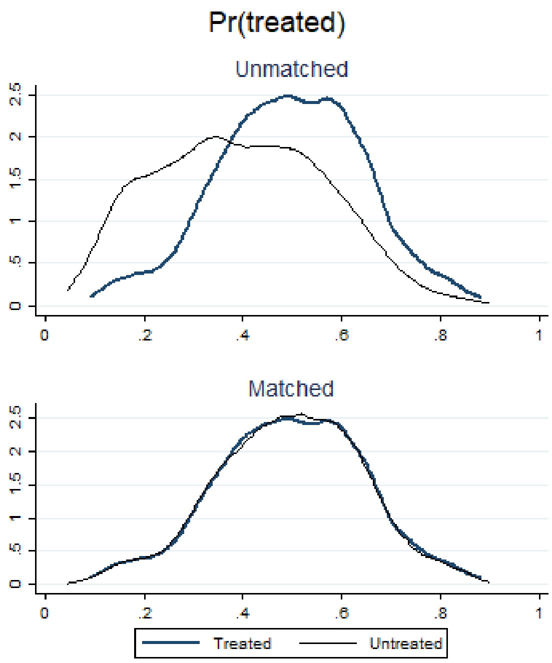
Figure A1 Propensity score distribution

Panel A: Total

Base specification

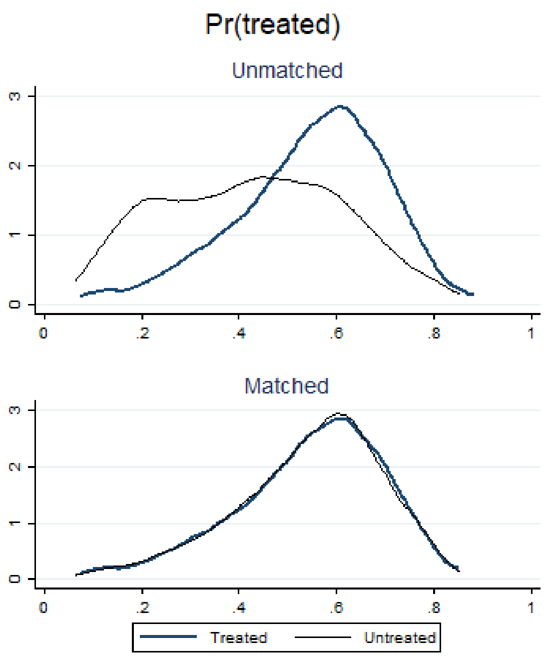


Full specification

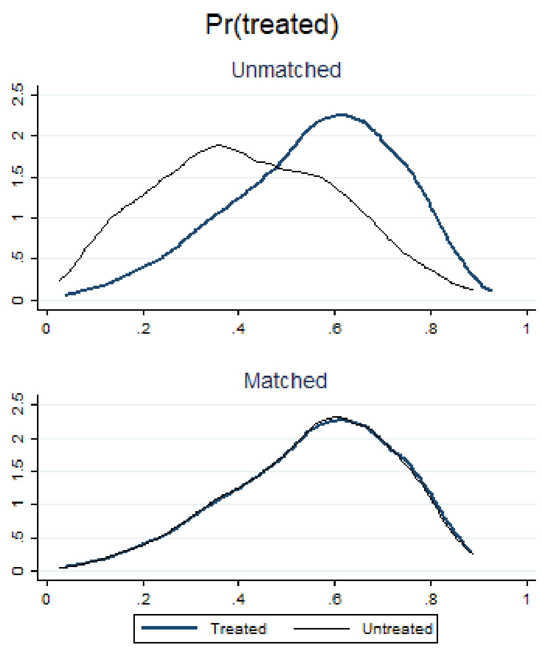


Panel B: 18-29 years old

Base specification

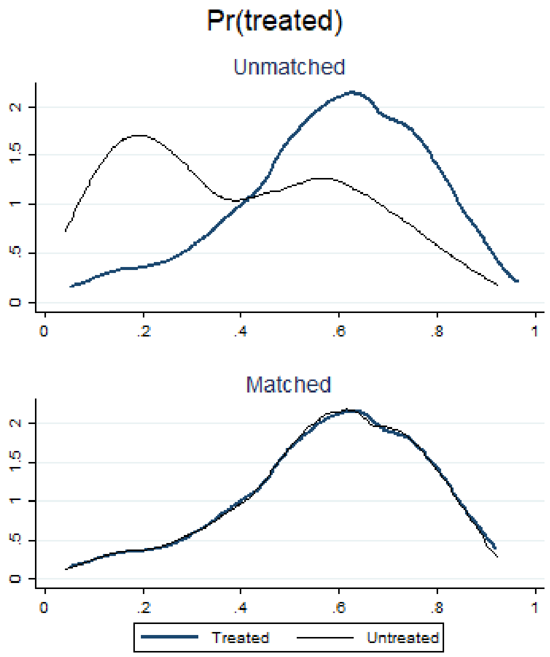


Full specification

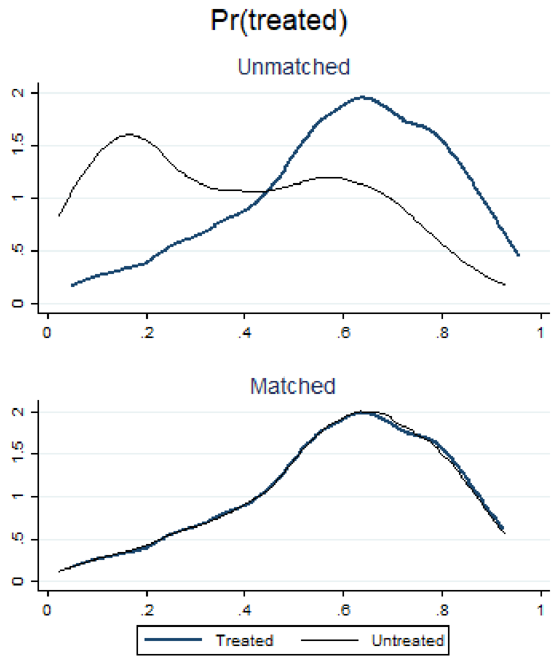


Panel C: 18-29 years old (ISCED 3-4)

Base specification

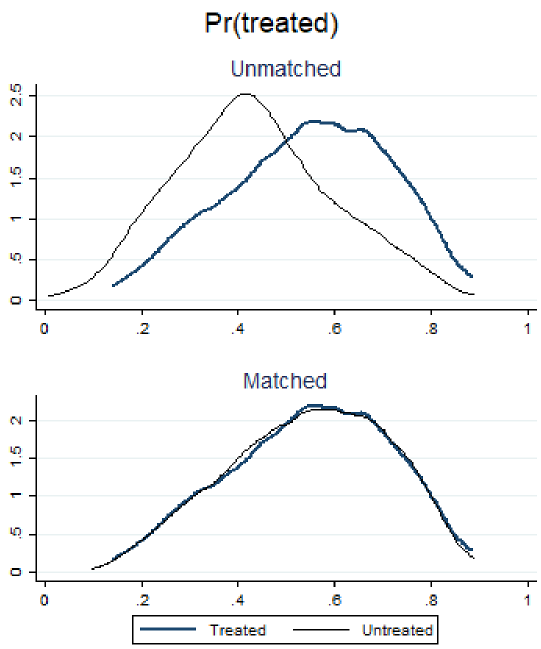


Full specification

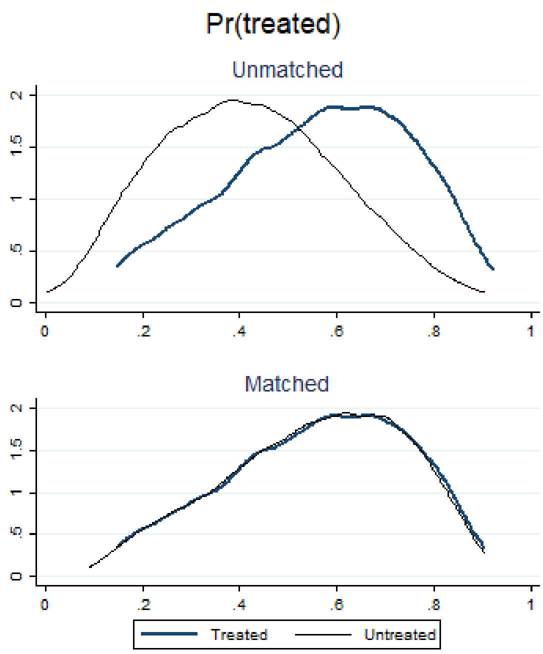


Panel D: 18-29 years old (ISCED 5-6)

Base specification



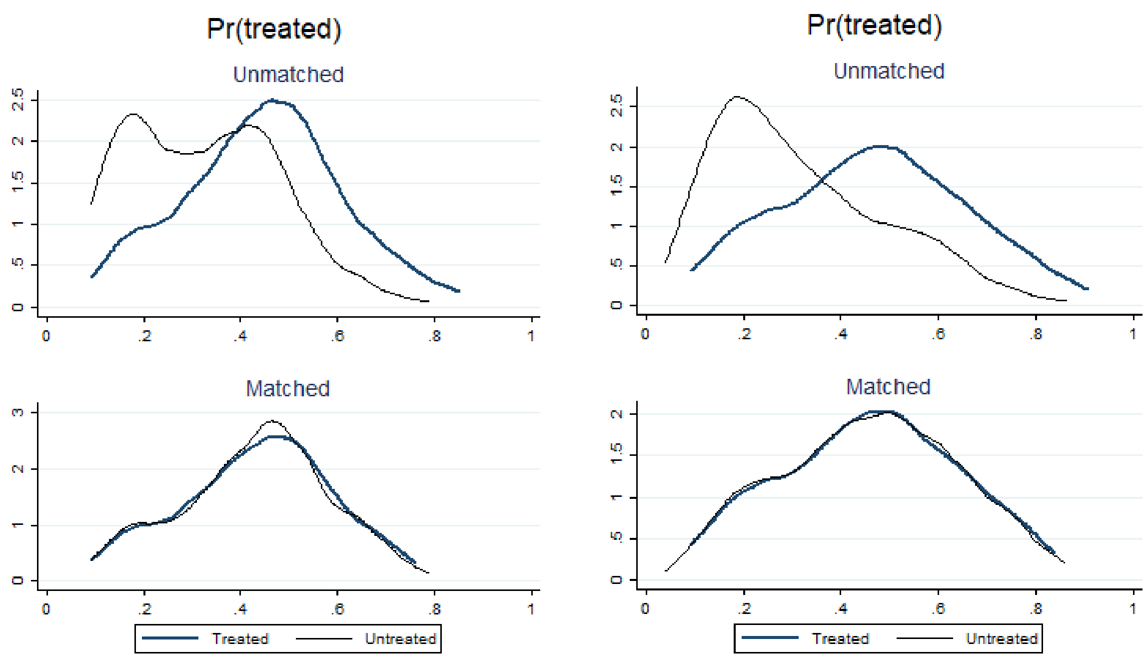
Full specification



Panel E: 30-68 years old

Base specification

Full specification



Notes: Presented are kernel densities (bandwith=0.06) of the propensity score before and after matching in base and full specifications. Base specification includes only socio-demographics and full specification includes socio-demographics and cognitive and non-cognitive skills.

instytut badań strukturalnych

www.ibs.org.pl