

Palczyńska, Marta

Article

Overeducation and wages: The role of cognitive skills and personality traits

Baltic Journal of Economics

Provided in Cooperation with:

Baltic International Centre for Economic Policy Studies (BICEPS), Riga

Suggested Citation: Palczyńska, Marta (2021) : Overeducation and wages: The role of cognitive skills and personality traits, Baltic Journal of Economics, ISSN 2334-4385, Taylor & Francis, London, Vol. 21, Iss. 1, pp. 85-111, <https://doi.org/10.1080/1406099X.2021.1950388>

This Version is available at:

<https://hdl.handle.net/10419/267589>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Overeducation and wages: the role of cognitive skills and personality traits

Marta Palczyńska

To cite this article: Marta Palczyńska (2021) Overeducation and wages: the role of cognitive skills and personality traits, *Baltic Journal of Economics*, 21:1, 85-111, DOI: [10.1080/1406099X.2021.1950388](https://doi.org/10.1080/1406099X.2021.1950388)

To link to this article: <https://doi.org/10.1080/1406099X.2021.1950388>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 13 Jul 2021.



Submit your article to this journal [↗](#)



Article views: 836



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Overeducation and wages: the role of cognitive skills and personality traits

Marta Palczyńska  ^{a,b}

^aInstitute for Structural Research (IBS), Warsaw, Poland; ^bSGH - Warsaw School of Economics, Warsaw, Poland

ABSTRACT

This article investigates the extent to which personality traits and cognitive skills can be seen as potential determinants of overeducation, and can explain the 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 18–29, but not among workers aged 30–68. Among younger workers, agreeable individuals are more likely to be overeducated, while conscientious individuals are less likely to be overeducated. Moreover, lower numeracy skills are associated with higher probability of being overeducated.

ARTICLE HISTORY

Received 19 November 2020
Accepted 28 June 2021

KEYWORDS

Overeducation; educational mismatch; wages; personality traits; cognitive skills; non-cognitive skills

JEL

D91; I26; J24; J31

1. Introduction

In the last three decades, many countries have experienced rapid improvements in human capital. After Poland transitioned from having a centrally planned economy to having a market economy, tertiary education expanded quickly in the country. The share of university graduates among the working-age population increased from 12% in 2000 to 28% in 2015. Among the youngest cohorts (25–34 years old), almost 43% had tertiary education in 2015.¹ On the one hand, this investment in human capital is viewed as positive for the economic development of the country. It has, for example, been argued that having a highly educated workforce would facilitate Eastern Europe's economic convergence with Western Europe (Caselli & Tenreyro, 2006). On the other hand, this rapid expansion of higher education raises concerns about the capacity of the economy to make good use of the growing highly skilled workforce.

Overeducation, which arises when a worker's qualifications are higher than the qualifications required by his or her job, can have serious consequences both for individuals and societies. The literature provides consistent evidence that overeducation is associated

CONTACT Marta Palczyńska  marta.palczynska@ibs.org.pl  Institute for Structural Research (IBS), Irysowa 18c, Warsaw 02-660, Poland

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

with wage penalties (Caroleo & Pastore, 2018; Korpi & Tåhlin, 2009; Montt, 2017; Rubb, 2003) and job dissatisfaction (Green & Zhu, 2010; Verhaest & Omey, 2006). From the perspective of policymakers, especially in countries where tertiary education is financed by the state, overeducation can be considered a waste of resources, as public investments in schooling are not being put to productive use. The issues surrounding the costs of overeducation and the potential policy responses to it are becoming even more relevant as the incidence of overeducation has increased or remained unchanged in almost all of EU countries over the 2001–2011 period (McGuinness et al., 2018). The upward trend in overeducation in Poland has been confirmed by Baran (2018) over the 2006–2014 period.

The aims of this article are to explore the relationships between cognitive and non-cognitive skills and the probability of being overeducated, and to investigate the role of these skills in explaining the overeducation wage penalty. While there is a large literature on the incidence and consequences for wages of overeducation, analyses accounting for the heterogeneity of skills among workers with the same level of education are sparse. Recently, the impact of non-cognitive skills on labour market outcomes has been increasingly recognised (for reviews, see Almlund et al., 2011; Borghans et al., 2008). Non-cognitive skills include a variety of character skills, such as personality traits, preferences, and motivations. Previous research has provided solid evidence that non-cognitive skills predict wages (e.g. Heckman & Kautz, 2012; Lindqvist & Vestman, 2011), but research on the relationship between non-cognitive skills and overeducation is limited, with the exceptions of Blázquez and Budría (2012), Kupets (2016), and Sohn (2010).

There are, however, reasons to expect that non-cognitive skills are related to the probability of being overeducated. An extended job search model (Caliendo et al., 2015) can potentially explain a mechanism behind this relationship. Unlike standard job search models in which unemployed individuals have perfect information about the payoffs of their search efforts, this model assumes that each individual has subjective beliefs about the potential payoffs of his or her job search that depend on the person's non-cognitive skills. If the skills negatively related to the probability of being overeducated are also productivity-enhancing, accounting for the heterogeneity of these skills among workers with the same level of education would partially explain the overeducation wage penalty.

Unobserved heterogeneity has been identified in the literature as one of the main challenges facing researchers attempting to causally estimate the overeducation wage penalty (Leuven & Oosterbeek, 2011). Workers with a given level of education may differ with respect to their skills, motivations, or on-the-job training. The main contribution of this paper is to rule out some of these usually unobserved factors: namely, workers' cognitive and non-cognitive skills. The second contribution is to explore the relationships between workers' cognitive and non-cognitive skills and their probability of being overeducated. 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 standardized achievement tests (collected in the Programme for International Assessment of Adult Competencies – PIAAC) that assess numeracy and problem-solving in technology-rich environments. Non-cognitive skills are measured by the Grit Scale and the Big Five framework,² which Kautz et al. (2014, p. 9) described as providing 'the longitude and latitude of non-cognitive skills, by which all more narrowly defined skills may be categorized'. Using the propensity score matching (PSM) approach

allows me to check whether the results obtained using linear regression are or are not attributable to common support problems.

The remainder of this paper is structured as follows. Section 2 provides a summary of the relevant theoretical and empirical literature. Section 3 describes the data used, the different measures of overeducation proposed in the literature, and the methodology applied. Section 4 presents the main results, and the findings of a sensitivity analysis that corroborate the results. Section 5 concludes.

2. Related literature

2.1. Theoretical explanations for overeducation

There are several competing labour market theories that seek to explain the differences in wages between workers. Human capital theory (Becker, 1964; Schultz, 1961) assumes that wages are solely determined by the productivity of workers. In its simplest version, it argues that workers' wages depend solely on their attained education, and that returns to years of overeducation should therefore be equal to the returns to years of required education. However, empirical research on the earnings effects of being mismatched provides robust evidence that overeducated individuals earn less than well-matched individuals with the same level of education. Many authors have pointed out that individuals with a given level of education have heterogeneous endowments of other productive characteristics, such as cognitive or non-cognitive skills, and that these differences in characteristics could help to explain the overeducation wage penalty (Chevalier, 2003; Green & 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 may be consistent with the human capital model. Thus, these penalties may actually be 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 solely by the characteristics of the job, and that a worker's education only determines his or her position in the queue for the best jobs. However, the empirical regularities are not explained by Thurow's job competition theory. Assignment models (for reviews, see Sattinger, 1993, 2012) are often seen as being better able to explain the wage consequences of educational mismatch. These models assume that wages are determined by the characteristics of both the job and the worker, and are a solution to the problem of the allocation of heterogeneous workers to heterogeneous jobs.

The career mobility model (Sicherman & Galor, 1990) suggests that workers might choose a position for which they are overeducated if the position offers them a higher probability of being promoted by providing on-the-job training and experience. This theory implies that overeducation is a temporary phenomenon. Matching theory (Jovanovic, 1979) also predicts that overeducation is temporary. This model posits that mismatch arises because of imperfect information about the quality of the match, and that with increasing tenure, the mismatch between a worker and an employer is detected, and the worker is able to improve the match through a job search.

This analysis relies on the human capital model, which defines human capital as all skills that contribute to production. Cognitive and non-cognitive skills are therefore

considered dimensions of human capital alongside traditional measures of human capital, such as years of schooling or experience. Thus, the analysis investigates whether accounting for the heterogeneity of these skills among workers with the same level of education can partially explain the overeducation wage penalty: i.e. whether the differences in wages between overeducated and well-matched individuals are generated by the differences in their productive characteristics other than education: namely, their cognitive and non-cognitive skills.

2.2. Empirical evidence on the overeducation wage penalty

The research on the wage consequences of overeducation has extensively used a modification of the Mincerian earnings equation introduced by Duncan and Hoffman (1981). The so-called ORU model decomposes actual years of education into the three categories of overeducation, required education, and undereducation. The results of empirical studies on the effects of overeducation on wages have been very consistent across the geographical locations and the time periods analysed: namely, that 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) estimated that the average rate of return to required years of education is 9.6%, the return to a year of overeducation is 5.2%, and the penalty for every year of undereducation is -4.8%. Another approach is to estimate the wage penalty for overeducation by comparing workers with the same level of education by means of a dummy variable (Verdugo & Verdugo, 1989). Across studies using a dummy variable, the average overeducation wage penalty has been estimated at 15.3% (McGuinness, 2006). Wincenciak (2016) applied comparable specifications to Poland, and obtained similar results: an overeducation wage penalty of 14.3% for the latter model; and values of 4.6%, 11%, and -5.6%, respectively, for the ORU specification.

There are different approaches to overcoming the problem of unobserved heterogeneity when estimating the overeducation wage penalty. First, several studies have used fixed effects (FE) models in order to address this problem. Bauer (2002) showed that the wage penalty becomes smaller when controlling for individual fixed effects. However, it was later demonstrated that when measurement error is accounted for, panel estimates of overeducation on wages are close to ordinary least squares (OLS) (Dolton & Silles, 2008; Verhaest & Omey, 2012). Moreover, the FE approach was criticized because the effects were identified from only a small fraction of individuals changing their educational match; and because the strict exogeneity assumption may not hold, as the unobserved characteristics of individuals are likely to change with their educational match status (Leuven & Oosterbeek, 2011). Second, instrumental variables are used to account for unobserved heterogeneity (Kleibrink, 2016; Korpi & Tåhlin, 2009). As Leuven and Oosterbeek (2011) pointed out, these instruments are weak, and might not fulfil exclusion restrictions. Finally, proxies of skills (Allen & Van der Velden, 2001; Chevalier, 2003; Green & 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 have so far been 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 hardly changed the returns to required

education, overeducation, and undereducation. However, the analysis included only a single dimension of personality: namely, locus of control.

While most of the existing analyses of the overeducation wage penalty used OLS, several studies used non-parametric methods. For example, McGuinness (2008) and McGuinness and Sloane (2011) found that PSM estimates of the overeducation wage penalty were close to OLS estimates based on samples of graduates from Northern Ireland and the UK, respectively. Similarly, Lamo and Messina (2010) used PSM to assess the impact of overeducation on wages in the Estonian working-age population. They also concluded that these estimates are in line with those generated by OLS. The evidence showing that OLS and PSM produce comparable results suggests that there is no common support problem. However, none of these studies investigated the role of personality traits as a confounding factor, and only McGuinness (2008) used an ability measure in his model.

2.3. Overeducation and personality traits

Non-cognitive skills might affect the probability of being overeducated in multiple ways. The most important channel seems to be the job search effort. An extended job search model (Caliendo et al., 2015) assumes that each individual has subjective beliefs about the potential payoffs of a job search that depend on individual's personality: namely, locus of control. Individuals with an internal (as opposed to an external) locus of control trust that life events depend on their own actions (Almlund et al., 2011). Therefore, these individuals can be expected to send out more job applications and to have a higher reservation wage. Empirical evidence supports the predictions of the model that the number of job applications and the reservation wage of each individual depends on his or her locus of control (Caliendo et al., 2015; McGee, 2015). These results might suggest that workers with a greater internal locus of control have a wider choice of job offers, and are less likely to accept a position with requirements below their qualification levels. Neuroticism – which is characterized by emotional instability, vulnerability to stress, and a lack of self-confidence (Almlund et al., 2011) – and having an external locus of control are strongly correlated. Moreover, these characteristics are indicators of a higher order trait: namely, core self-evaluation (Judge et al., 2005, 2002). Thus, it seems reasonable to expect that neurotic individuals are more likely to be overeducated. Other personality traits may also be positively related to individuals' job search efforts, including conscientiousness, which is the tendency to be organized, responsible, and hardworking; and grit, which is defined as perseverance and passion in pursuing long-term goals. Therefore, I assume that there is a negative relationship between these traits and the probability of being overeducated.

Other potential mechanisms linking personality traits and overeducation include negotiations and social networks. Agreeable individuals who have the tendency to act in a cooperative, unselfish manner are more likely than their less agreeable counterparts to accept unfair offers (Nguyen et al., 2011). Therefore, it seems reasonable to assume that agreeableness is positively related to being overeducated, as agreeable individuals are more prone to accept job offers below their qualification levels. Extraverts typically have broader social networks (Heineck & Anger, 2010), which could be helpful in finding a job that matches their qualifications. Open individuals are typically curious, imaginative, and have a wide range of interests. The relationship between openness and

being overeducated is ambiguous. In line with these conceptual considerations, I expect to find that neuroticism and agreeableness are positively associated with being overeducated; while conscientiousness, extraversion, and grit are negatively associated with being overeducated. Some of these traits are also related to wages. For Poland, it has been shown that conscientiousness is positively associated with wages, while neuroticism and agreeableness are negatively associated with wages (Palczyńska & Świst, 2018).

To the best of my knowledge, there are only three economic papers on the relationship between personality traits and overeducation. The first of these studies used panel data from Germany to show that conscientiousness, extraversion, and having an external locus of control decrease the probability of remaining overeducated, while openness increases it (Blázquez & Budría, 2012). The second study showed that in the United States, the locus of control is not related to the probability of being overeducated (Sohn, 2010). The third study analysed four non-EU transition economies (Armenia, Georgia, Macedonia, and Ukraine), and showed that conscientiousness and grit are not related to the probability of being overeducated, with the exception of an unexpected finding of a positive relationship between conscientiousness and overeducation in Ukraine and in the pooled sample of countries (Kupets, 2016).

3. Data and methods

3.1. Data

This study uses 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, which was coordinated by the OECD and conducted in 2011/2012, with information from 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 that was not available in the international study. The postPIAAC sample is made up of PIAAC respondents who were living in Poland during the postPIAAC fieldwork. The sample size is 5224 completed interviews with individuals who were among the 9366 Polish participants in PIAAC in 2011/2012. The data collection for postPIAAC was carried out from October 2014 to February 2015. This implies that the interval between the interviews for each individual respondent ranged from 2.5 to 3.5 years.

A unique feature of the PIAAC dataset is that it includes direct measures of cognitive skills. The respondents performed exercises in literacy, numeracy, and problem-solving in technology-rich environments. As the last domain was designed only for people with computer experience, 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 the computer; below level 1, level 1, and level 2/level 3. The postPIAAC study contains self-report scales of personality from the Big Five Inventory– Short (BFI-S) (Gerlitz & Schupp, 2005; John et al., 1991) and the Short Grit Scale (Grit-S) (Duckworth & Quinn, 2009), which are not available in the PIAAC study. The analysis uses the factor scores from the six-factor oblique model, with an additional factor loading the reverse-worded items for the Big Five questions, as it is the best-fitted model obtained in the psychometric analysis of the scale (Palczyńska & Świst, 2016); and the factor scores from the

unidimensional model for Grit-S, as Grit-S is often used as a single factor (Eskreis-Winkler et al., 2014). The scores for the BFI-S subscales and the Grit-S are standardized with a mean of zero and a standard deviation of one within the total working-age population.

The analyses are restricted to people in paid employment (excluding the self-employed) who answered all the relevant questions. The top and the bottom 1% of the observations with respect to working hours are removed. 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 included in the analysis, and are treated as well-matched. The overrepresentation of people aged 19–26 in the PIAAC sample allows for a separate analysis for younger (aged 18–29) and older (aged 30–68) workers. Younger workers face different constraints in the labour market, and have a higher risk of persistent overeducation (Kiersztyn, 2013). The sample size does not allow for a more detailed heterogeneity analysis with respect to age.

3.2. Educational mismatch measures

The empirical application of the educational mismatch concept requires a measure of the required level of education for a job. In the literature, three ways of measuring the required level of education have been proposed. First, some studies rely on the evaluation of each occupation by a professional job analyst (JA). This provides an exogenous measure of the required education. However, such evaluations are available for a limited number of countries only, and are updated infrequently. The second method uses the self-assessment (SA) of workers about the level of education required to do or, alternatively, to get their job. This measure allows for heterogeneous education requirements within the same occupation. The third measure is derived from realized job matches (RM) (Verdugo & Verdugo, 1989). For each occupation, the required education level is defined as the mean of the completed years of schooling of all workers in the same occupation. A worker is considered to be overqualified if his or her level of education exceeds the mean in his or her occupation by at least one standard deviation. Kiker et al. (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 research on overeducation has been conducted for more than three decades, there is no agreement in the literature about how required education should be measured.³ The measures used in these analyses were chosen mainly based on their availability, and not on their properties. Many of these studies examined whether the results of the earnings equation were sensitive to the measure used, and concluded that the results were independent of the measure of required education used (Chiswick & Miller, 2010; Cohn & Khan, 1995; Groot & van den Brink, 2000; Kiker et al., 1997). Only a few of these studies attempted to validate these measures. van der Velden and van Smoorenburg (2000) evaluated a SA and a JA measure within one wage equation, and concluded that the self-assessment method gave more accurate estimates. Verhaest and Omey (2006) compared JA, RM, and two types of SA measures using the same method of evaluation, but did not report that one of these methods of measurement was superior to the others. According to Hartog (2000), JA is ‘conceptually superior’, but its measurement is often flawed. Therefore, there are no clear guidelines in the literature regarding the appropriate measure of required education.

This analysis uses workers' self-assessments 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 those in the question on the highest completed level of education. The answers to both questions were converted into years of education (the correspondence table can be found in Table A1 in the appendix). These two questions combined give the indicator of being overeducated:

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

where EDU_i is individual i 's years of completed education, and REQ_i is the required years of completed education in individual i 's job.

3.3. Method

In this paper, I examine the role of cognitive and non-cognitive skills in the overeducation wage penalty by comparing the estimates from two specifications: without and with these characteristics. In terms of the methodological approach, I use matching methods. Additionally, I also estimate the overeducation wage penalty using linear regression, for the sake of comparison with the previous literature that used this method.

I use matching estimation based on the propensity score (Rosenbaum & 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 the unconfoundedness or conditional independence assumption (CIA), which states that conditional on observable variables that influence selection into treatment, the treatment status is assumed to be randomized. The parameter of interest in the analysis is the average treatment effect on the treated (ATT). In the available data, the non-cognitive skills measures are contemporaneous to the outcomes. Thus, the data do not allow me to rule out the reverse causality problem in the selection model. However, the causal interpretation of the results is supported by previous evidence showing that personality traits are stable over time (Cobb-Clark & Schurer, 2012), and that a major work-related life event, such as an involuntary job loss, tends to have a very limited or temporary impact on personality (Anger et al., 2017; Preuss & Hennecke, 2018).

To empirically investigate whether often unobserved variables such as cognitive skills and personality traits affect estimators based on the CIA assumption, I follow an approach analogous to that of Caliendo et al. (2017), who assessed the impact of often unobserved characteristics on the estimated effectiveness of active labour market programmes in Germany. I compare the results when these characteristics are included in the model (full specification), and when they are not included in the model (base specification).

I estimate the propensity score using a logit model:

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

where $OVER_i$ is the treatment indicator for working in a job for which one is overeducated, and X_i is a vector of individual characteristics. The control group ($OVER_i = 0$) consists of well-

matched individuals ($EDU_i = REQ_i$). I also check the sensitivity of the results when undereducated individuals ($EDU_i < REQ_i$) are included in the control group (with well-matched individuals). In the base specification, the vector of individual characteristics includes the following: age, age squared, gender, level of education, field of education, a dummy for still in education, experience, experience squared, a dummy for living with a partner, a dummy for having children aged 0–6, mother's level of education, number of books in the home when the respondent was age 16, and the registered unemployment rate in 2014 (on the county level). In the full specification, the individual characteristics additionally include numeracy skills, ICT skills level, and personality traits. These equations are run for the whole sample, and separately by age groups (18–29 and 30–68) and by education levels: upper secondary and post-secondary non-tertiary (ISCED 3–4, according to ISCED-97⁴) and tertiary (ISCED 5–6, according to ISCED-97) in the younger age group. For the base and the full specifications for tertiary graduates (ISCED 5–6), a dummy for attending a private university is added. Job characteristics are not included in the model, as they are most likely intermediate outcomes. Thus, if job characteristics were included in the model, a downward bias would likely influence the overall effects of overeducation on wages. The differences in the mean characteristics between overeducated and matched individuals are shown in Table A2 in the appendix.

I consider two matching algorithms with different parameters: nearest neighbour (NN) matching (one and four 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:

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

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

Finally, I also estimate a wage equation using a modification of Mincerian wage equation with a dummy for overeducation⁵:

$$\ln w_i = \theta X_i + \beta OVER_i + u_i,$$

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. The vector of individual characteristics includes the same variables as the aforementioned PSM model.

3.4. Educational mismatch incidence

In the analysed data, one-third of Polish dependent workers are overeducated, while 14% are undereducated (Table 1). Overeducation is more common among workers under age 30 (43%), and in this group, tertiary graduates are more likely to be overeducated (47%).

Table 1. Incidence of the educational mismatch.

	Total	Age 18–29	Age 30–68	Age 18–29		Age 30–68	
				ISCED ^a 3–4	ISCED 5–6	ISCED 3–4	ISCED 5–6
Undereducated	14%	11%	15%	15%	2%	21%	1%
Matched	53%	47%	55%	43%	51%	49%	65%
Overeducated	33%	43%	31%	42%	47%	30%	34%
Observations	2661	1702	959	863	758	527	389

Notes: ^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012). Sample weights are used in the analysis.

Source: Own calculations based on postPIAAC data.

Wincenciak (2016), who used the RM approach to measure educational mismatch in Poland in 2013, obtained lower estimates of overeducation incidence (25%) and higher estimates of undereducation incidence (25%).

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 are found to be overeducated (Table 2). A larger share of the younger individuals aged 18–29 (48%) than of the older individuals aged 30–68 (36%) are shown to be overeducated. The educational level is not found to be associated with the incidence of overeducation.

4. Results

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

The effects of personality on selection into overeducation among younger workers are driven by the tertiary education graduates (ISCED 5–6) (Table 3). As expected, conscientiousness decreases the risk of being overeducated, while agreeableness is associated with a higher risk of overeducation. These traits are not related to being overeducated among people aged 30–68. A potential explanation for the differences observed between the age groups is that these traits are expected to be important when entering the labour market: i.e. when applying for a job and negotiating wages. Relationships between the risk of overeducation and the other personality traits are not observed. In

Table 2. Incidence of overeducation in a reduced sample.

	Total	Age 18–29	Age 30–68	Age 18–29		Age 30–68	
				ISCED ^a 3–4	ISCED 5–6	ISCED 3–4	ISCED 5–6
Matched	61%	52%	64%	51%	51%	63%	63%
Overeducated	39%	48%	36%	49%	49%	37%	37%
Observations	1598	1091	507	512	548	255	240

Notes: ^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012). Sample weights are used in the analysis.

Source: Own calculations based on postPIAAC data.

Table 3. Selection into overeducation.

	Total	Age 18–29	Age 18–29 (ISCED ^a 3–4)	Age 18–29 (ISCED 5–6)	Age 30–68
Numeracy	–0.234* (0.108)	–0.042 (0.104)	–0.169 (0.162)	0.052 (0.153)	–0.263 (0.153)
ICT (no exp. or failed test)	ref.	ref.	ref.	ref.	ref.
ICT (refused test)	0.119 (0.399)	1.059* (0.465)	1.308* (0.656)	0.852 (0.688)	0.006 (0.466)
ICT (below level 1)	–0.433 (0.424)	0.569 (0.403)	0.478 (0.570)	0.664 (0.865)	–0.551 (0.535)
ICT (level 1)	–0.582 (0.346)	0.481 (0.345)	0.438 (0.444)	0.392 (0.713)	–0.795 (0.455)
ICT (level 2 and 3)	0.028 (0.367)	0.422 (0.336)	0.207 (0.422)	0.423 (0.649)	0.185 (0.483)
B5: Conscientiousness	0.055 (0.222)	–0.458* (0.209)	–0.105 (0.377)	–0.697* (0.286)	0.322 (0.336)
B5: Extraversion	–0.122 (0.143)	–0.163 (0.145)	–0.057 (0.233)	–0.142 (0.168)	–0.140 (0.178)
B5: Agreeableness	0.025 (0.215)	0.577* (0.236)	0.208 (0.397)	0.878** (0.313)	–0.198 (0.326)
B5: Openness	–0.107 (0.166)	–0.143 (0.158)	–0.269 (0.256)	–0.179 (0.190)	–0.151 (0.223)
B5: Neuroticism	0.095 (0.105)	–0.002 (0.080)	0.117 (0.131)	–0.105 (0.138)	0.135 (0.148)
Grit	–0.057 (0.100)	–0.084 (0.086)	–0.067 (0.134)	–0.081 (0.122)	–0.021 (0.138)
Observations	1598	1091	512	548	507

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 for still in education, experience, experience squared, a dummy for living with a partner, a dummy for having children aged 0–6, mother's level of education and number of books in the home when the respondent was age 16, local unemployment rate; in the 18–29 (ISCED 5–6) subpopulation, additionally a dummy for attending a private university. Numeracy and non-cognitive skills are standardized.

^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012). Full estimation results are included in Table 8 in the appendix.

addition, 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 the computer or who failed the test in the 18–29 age group. This finding might be related to the tendency to avoid competition, rather than to the ICT skills level itself. Lower numeracy skills are associated with being overeducated, and this result is mainly driven by workers aged 30–68.

The overall measures of covariate imbalance suggest that the kernel matching algorithm with the bandwidth of 0.06 provides the best balance of covariates after matching (appendix Table A4). However, there is a problem when attempting to find a good balance between treated and controls in the four models: the base and the full specification for the total population and for the 18–29 age group (ISCED 3–4). For these models, the mean standardised bias is in the 5%–7% range, and thus is slightly above the 5% threshold for empirical studies (Caliendo & Kopeinig, 2008). Statistics on the quality of the matching for all of the algorithms used are reported in Table A4, and the distributions of propensity scores before and after matching are displayed in Figure A1 in the appendix.

The inclusion of cognitive skills and personality traits leads to only a marginal or no reduction in the overeducation wage penalty (comparing the full specification and the base specification) (Table 4). The reduction is the biggest for the 18–29-year-olds with secondary education. For this group, the overeducation penalty is reduced from 15.1% to

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

	PSM			OLS			N
	Base	Full	Difference	Base	Full	Difference	
Total	-0.163*** (0.022)	-0.168*** (0.023)	0.0048 (0.0083)	-0.141*** (0.033)	-0.134*** (0.032)	-0.0068 (0.8816 ^a)	1598
Age 18–29	-0.144*** (0.033)	-0.139*** (0.028)	-0.0050 (0.0131)	-0.132*** (0.032)	-0.128*** (0.032)	-0.0044 (0.5340)	1091
Age 18–29 (ISCED ^b 3–4)	-0.128*** (0.038)	-0.103** (0.037)	-0.0246 (0.0288)	-0.151** (0.054)	-0.133* (0.054)	-0.0185 (3.0332)	512
Age 18–29 (ISCED 5–6)	-0.175*** (0.030)	-0.169*** (0.038)	-0.0068 (0.0265)	-0.155*** (0.035)	-0.157*** (0.035)	0.0019 (0.0350)	548
Age 30–68	-0.129* (0.052)	-0.148* (0.057)	0.0193 (0.0251)	-0.137** (0.042)	-0.135** (0.042)	-0.0027 (0.0502)	507

Notes: PSM is ATT kernel matching estimator (bandwidth=0.06). In italics ATT with MSB>5%. Standard errors are in parentheses, and are based on bootstrapping with 50 replications. Standard errors for the differences in ATTs are based on bootstrapped robust Hausman tests with 50 replications (see Cameron & Trivedi, 2009, for details) ^aF statistic for the hypothesis that the coefficients are equal to each other.

^bInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

***/**/*Indicate statistical significance at the 0.1%/1%/5%-level.

13.3% in the OLS model, and from 12.8% to 10.3% in the PSM model, but none of the differences are statistically significant. When individuals with different educational levels are analysed separately, the effects are more heterogenous in the PSM model than in the OLS model. In the PSM model, the overeducation wage penalty is 10% for individuals with secondary education and 17% for tertiary graduates; while in the OLS model, the corresponding values are 13% and 16%.

As a sensitivity analysis, I estimate models in which undereducated individuals are treated as matched, and are included in the control group (Table 5). The main results described above are shown to hold, as the inclusion of cognitive skills and personality

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

	PSM			OLS			N
	Base	Full	Difference	Base	Full	Difference	
Total	-0.167*** (0.023)	-0.168*** (0.024)	0.0012 (0.0084)	-0.149*** (0.031)	-0.140*** (0.030)	-0.0080 (1.4696 ^a)	1818
Age 18–29	-0.150*** (0.022)	-0.146*** (0.026)	-0.0048 (0.0106)	-0.126*** (0.032)	-0.123*** (0.032)	-0.0029 (0.2909)	1220
Age 18–29 (ISCED ^b 3–4)	-0.134** (0.047)	-0.139** (0.041)	0.0050 (0.0272)	-0.141** (0.051)	-0.131* (0.052)	-0.0101 (1.4458)	608
Age 18–29 (ISCED 5–6)	-0.177*** (0.039)	-0.165*** (0.036)	-0.0118 (0.0246)	-0.154*** (0.035)	-0.154*** (0.036)	-0.0003 (0.0007)	558
Age 30–68	-0.134** (0.050)	-0.155** (0.048)	0.0208 (0.0308)	-0.151*** (0.039)	-0.143*** (0.039)	-0.0074 (0.4527)	598

Notes: PSM is ATT kernel matching estimator (bandwidth = 0.06). In italics ATT with MSB>5%. Standard errors are in parentheses and based on bootstrapping with 50 replications. Standard errors for the differences in ATTs are based on bootstrapped robust Hausman tests with 50 replications (see Cameron & Trivedi, 2009, for details) ^aF statistic for the hypothesis that the coefficients are equal to each other.

^bInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

***/**/*Indicate statistical significance at the 0.1%/1%/5%-level.

traits leads to no reduction in the overeducation wage penalty. When undereducated individuals are included in the control group, the estimates of the overeducation wage penalty are slightly higher. This is to be expected, as undereducated individuals earn more than matched individuals with the same level of education. Thus, including undereducated individuals in the control group increases the average wage in the control group.

5. Discussion and conclusions

The aim of this paper was to investigate whether heterogeneity in cognitive and non-cognitive skills explains the likelihood of being overeducated and the overeducation wage penalty. Thus, it has extended research on the consequences of overeducation by showing that 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 are suspected to be the source of omitted variable bias. However, the results of this analysis suggest that matching inefficiencies, and not differences in these dimensions of human capital, account for the overeducation wage penalty.

Some personality traits – namely, agreeableness and conscientiousness – were found to be related to the probability of being overeducated among workers aged 18–29; but no personality traits were shown to be related to being overeducated among workers aged 30–68. The reason for this difference between the age groups might be that these traits are probably important at the job application stage, when workers are entering the labour market, but with time, workers manage to improve the match between their position and their qualifications within the company, or by using professional networks they have built with experience. As expected, agreeableness was shown to be positively related to overeducation, while conscientiousness was found to be negatively related to overeducation. These findings are consistent with earlier results reported by Blázquez and Budría (2012) for conscientiousness, though they found agreeableness to be unrelated to overeducation. Lower cognitive skills were shown to be associated with being overeducated, and this result was mainly driven by workers aged 30–68.

Even though these skills were found to be related to being overeducated, the overall impact of including them in the model estimating overeducation wage penalty was insignificant. This might be because only some of the analysed skills were found to be related to being overeducated, and only in some subpopulations. Thus, the overall effects appeared to be limited. It is also possible that these skills are related to overeducation in occupations or industries in which wages are rigid, and are not related to cognitive and non-cognitive skills.

The analysis also uncovered heterogeneity in the overeducation wage penalty among educational groups, with overeducated workers with tertiary education experiencing a higher penalty than overeducated workers with secondary education. Similar results were obtained for the skills match in Poland by Liwiński and Pastore (2021), who showed that the quality of the skills-job match is related to the wages of tertiary education graduates, but not to the wages of secondary education graduates. This pattern may be attributable to the differences in the education premium between these levels of education. The differences in wages are bigger between workers with tertiary and secondary education than they are between workers with secondary and lower than secondary education (Strawiński et al., 2018).

The results also indicated that the differences between OLS and PSM estimates are quite substantial for some subpopulations, and are even in the opposite directions: i.e. for workers with secondary education, the PSM estimate of the overeducation wage penalty was found to be smaller than the OLS estimate, while for workers with tertiary education the PSM estimate was shown to be higher than the OLS estimate.

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

The measure of overeducation used in the analysis was based on workers' self-assessments. If the bias in the self-assessments was systematically related to the levels of cognitive skills, this could explain the observed relationship between overeducation and cognitive skills. If workers with lower cognitive skills have a greater tendency to systematically understate the requirements of their jobs than individuals who have the same level of education, and are doing the same job, but with higher skills, overeducated individuals would appear to have lower cognitive skills, but only because of the differences in the self-assessments of educational requirements. In my opinion, this is very unlikely. First, the phrasing of the requirements question used in the analysis limits this problem. As the respondents were asked about the recruitment standards – i.e. about the usual qualifications for getting this type of job – and not about the qualifications needed to perform the job, the job requirements should be more objective, and easier for the respondents to assess. Second, there are no conceptual considerations suggesting that workers with lower cognitive skills would understate the educational requirements of their job, the opposite tendency seems more probable.

Thus, the results of the analysis suggest that the heterogeneity in cognitive and non-cognitive skills analysed does not explain the differences in wages between overeducated and matched workers. 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 & Vera-Toscano, 2017). This paper has ruled out some of these potential sources: namely, basic cognitive skills and personality traits, as measured by the Big Five and the Grit Scale. The analysis did not include all of the potential sources of heterogeneity among workers. For example, motivations or vocational skills may also be relevant for the consequences of overeducation. Further research on this issue is needed.

Notes

1. Eurostat Labour Force Survey (LFS) data.
2. The Big Five model defines personality on five dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (the opposite of emotional stability) (McCrae & Costa, 1999). Grit is defined as perseverance and a passion for pursuing long-term goals (Duckworth et al., 2007).
3. For a detailed discussion of the advantages and disadvantages of the different measures, see: e.g., Hartog (2000) and Leuven and Oosterbeek (2011).

4. International Standard Classification of Education 1997 (UNESCO Institute for Statistics 2012).
5. The choice not to include undereducated individuals and 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 (Table A6 in the appendix).

Acknowledgements

I would like to thank two anonymous referees for their useful comments and suggestions. I also wish to thank Jan Baran and Iga Magda for their insightful comments, and Jan Gromadzki for calculating the LFS statistics. I also want to thank the participants of the 2016 EALE conference and Warsaw International Economic Meeting 2019 for their discussion and remarks. This publication was created as part of the 'Youth Employment PartnerSHIP' project, financed by Iceland, Liechtenstein and Norway through the EEA and Norway Grants Fund for Youth Employment, which seeks to combat unemployment among young people. It measures the effectiveness of instruments to support young people who are not working or studying in four countries: Poland, Spain, Italy, and Hungary.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the EEA and Norway Grants Fund for Youth Employment under grant number: 2017-1-008.

Notes on contributor

Marta Palczyńska is an economist at the Institute for Structural Research (IBS, Warsaw) and a PhD candidate at the Warsaw School of Economics. Her work focuses on the role of cognitive and non-cognitive skills in the labour market, educational mismatches, and the evaluation of the youth labour market policies. Previously, she worked at the Educational Research Institute (IBE, Warsaw), where she was involved in research projects on the situations of youth in the labour market (EXCEPT), and on the assessment of adults competences (PIAAC). She also cooperated with the OECD in the field of non-cognitive skills measurement.

ORCID

Marta Palczyńska  <http://orcid.org/0000-0003-3262-5745>

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(3), 434–452. <https://doi.org/10.1093/oep/53.3.434>
- Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S. J. Machin, & L. Woessmann (Eds.), *Handbook of economics of education* (Vol. 4, pp. 1–181). Elsevier BV.
- Anger, S., Camehl, G., & Peter, F. (2017). Involuntary job loss and changes in personality traits. *Journal of Economic Psychology*, 60, 71–91. <https://doi.org/10.1016/j.joep.2017.01.007>

- 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(18)(2), 41–63. <https://doi.org/10.18559/ebr.2018.2.3>
- Bauer, T. K. (2002). Educational mismatch and wages: A panel analysis. *Economics of Education Review*, 21(3), 221–229. [https://doi.org/10.1016/S0272-7757\(01\)00004-8](https://doi.org/10.1016/S0272-7757(01)00004-8)
- 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(3), 260–283. <https://doi.org/10.1080/09645292.2012.679338>
- Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 972–1059. <https://doi.org/10.3368/jhr.43.4.972>
- Caliendo, M., Cobb-Clark, D. A., & Uhlendorff, A. (2015). Locus of control and job search strategies. *Review of Economics and Statistics*, 97(1), 88–103. https://doi.org/10.1162/REST_a_00459
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- 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. <https://doi.org/10.1016/j.labeco.2017.02.001>
- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using stata*. Stata Press.
- Caroleo, F. E., & Pastore, F. (2018). Overeducation at a glance. Determinants and wage effects of the educational mismatch based on AlmaLaurea data. *Social Indicators Research*, 137(3), 999–1032. <https://doi.org/10.1007/s11205-017-1641-1>
- Caselli, F., & Tenreiro, S. (2006). Is Poland the next Spain? In R. H. Clarida, J. Frankel, F. Giavazzi, & K. D. West (Eds.), *NBER international seminar on macroeconomics 2004* (pp. 459–533). MIT Press.
- Chevalier, A. (2003). Measuring over-education. *Economica*, 70(279), 509–531. <https://doi.org/10.1111/1468-0335.t01-1-00296>
- 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(6), 1076–1085. <https://doi.org/10.1016/j.econedurev.2010.06.001>
- Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1), 11–15. <https://doi.org/10.1016/j.econlet.2011.11.015>
- Cohn, E., & Khan, S. P. (1995). The wage effects of overschooling revisited. *Labour Economics*, 2(1), 67–76. [https://doi.org/10.1016/0927-5371\(95\)80008-L](https://doi.org/10.1016/0927-5371(95)80008-L)
- de Grip, A., Bosma, H., Willems, D., & van Boxtel, M. (2008). Job-worker mismatch and cognitive decline. *Oxford Economic Papers*, 60(2), 237–253. <https://doi.org/10.1093/oep/gpm023>
- Dolton, P. J., & Silles, M. A. (2008). The effects of over-education on earnings in the graduate labour market. *Economics of Education Review*, 27(2), 125–139. <https://doi.org/10.1016/j.econedurev.2006.08.008>
- 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(6), 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of Personality Assessment*, 91(2), 166–174. <https://doi.org/10.1080/00223890802634290>
- Duncan, G. J., & Hoffman, S. D. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1), 75–86. [https://doi.org/10.1016/0272-7757\(81\)90028-5](https://doi.org/10.1016/0272-7757(81)90028-5)
- Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, 5(36). <https://doi.org/10.3389/fpsyg.2014.00036>
- Gerlitz, J.-Y., & Schupp, J. (2005). *Zur Erhebung der Big-Five-basierten persönlichkeitsmerkmale 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(4), 427–439. <https://doi.org/10.1080/00036840500427700>
- Green, F., & Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, 62(4), 740–763. <https://doi.org/10.1093/oep/gpq002>

- Groot, W., & van den Brink, H. M. (2000). Overeducation in the labor market: A meta-analysis. *Economics of Education Review*, 19(2), 149–158. [https://doi.org/10.1016/S0272-7757\(99\)00057-6](https://doi.org/10.1016/S0272-7757(99)00057-6)
- Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2), 131–147. [https://doi.org/10.1016/S0272-7757\(99\)00050-3](https://doi.org/10.1016/S0272-7757(99)00050-3)
- Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451–464. <https://doi.org/10.1016/j.labeco.2012.05.014>
- Heineck, G., & Anger, S. (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, 17(3), 535–546. <https://doi.org/10.1016/j.labeco.2009.06.001>
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory—Versions 4a and 5a*. 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(5, Part 1), 972–990. <https://doi.org/10.1086/260808>
- 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(2), 257–268. <https://doi.org/10.1037/0021-9010.90.2.257>
- 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(3), 693–710. <https://doi.org/10.1037/0022-3514.83.3.693>
- 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. <https://doi.org/10.1016/j.econedurev.2012.09.009>
- Kiker, B. F., Santos, M. C., & de Oliveira, M. M. (1997). Overeducation and undereducation: Evidence for Portugal. *Economics of Education Review*, 16(2), 111–125. [https://doi.org/10.1016/S0272-7757\(96\)00040-4](https://doi.org/10.1016/S0272-7757(96)00040-4)
- Kleibrink, J. (2016). Inept or badly matched? — Effects of educational mismatch in the labor market. *Labour (committee. on Canadian Labour History)*, 30(1), 88–108. <https://doi.org/10.1111/lab.12065>
- Korpi, T., & Tählin, M. (2009). Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics*, 16(2), 183–193. <https://doi.org/10.1016/j.labeco.2008.08.004>
- Kupets, O. (2016). Education in transition and Job mismatch: Evidence from the skills Survey in Non-EU transition economies. *The Journal of Comparative Economic Studies*, 11, 117–150. http://www.ces.kier.kyoto-u.ac.jp/jces_11.html
- 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(6), 1086–1099. <https://doi.org/10.1016/j.econedurev.2010.06.002>
- Leuven, E., & Oosterbeek, H. (Eds.), (2011). Overeducation and mismatch in the labor market. In E. A. Hanushek, S. Machin, & L. Woessmann, *Handbook of the Economics of education* (Vol. 4, pp. 283–326). Elsevier.
- Levels, M., van der Velden, R., & Allen, J. (2014). Educational mismatches and skills: New empirical tests of old hypotheses. *Oxford Economic Papers*, 66(4), 959–982. <https://doi.org/10.1093/oep/gpu024>
- Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 3(1), 101–128. <https://doi.org/10.1257/app.3.1.101>
- Liwiński, J., & Pastore, F. (2021). Are school-provided skills useful at work? Results of the wiles test. *Research in Higher Education*, 62, 72–97. <https://doi.org/10.1007/s11162-019-09569-5>
- McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (2, pp. 139–153). The Guilford Press.
- McGee, A. D. (2015). How the perception of control influences unemployed job search. *Industrial & Labor Relations Review*, 68(1), 184–211. <https://doi.org/10.1177/0019793914556245>

- McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20(3), 387–418. <https://doi.org/10.1111/j.0950-0804.2006.00284.x>
- McGuinness, S. (2008). How biased are the estimated wage impacts of overeducation? A propensity score matching approach. *Applied Economics Letters*, 15(2), 145–149. <https://doi.org/10.1080/13504850600721999>
- McGuinness, S., Bergin, A., & Whelan, A. (2018). Overeducation in Europe: Trends, convergence, and drivers. *Oxford Economic Papers*, 70(4), 994–1015. <https://doi.org/10.1093/oepp/gpy022>
- McGuinness, S., & Sloane, P. J. (2011). Labour market mismatch among UK graduates: An analysis using REFLEX data. *Economics of Education Review*, 30(1), 130–145. <https://doi.org/10.1016/j.econedurev.2010.07.006>
- Meroni, E. C., & Vera-Toscano, E. (2017). The persistence of overeducation among recent graduates. *Labour Economics*, 48, 120–143. <https://doi.org/10.1016/j.labeco.2017.07.002>
- Montt, G. (2017). Field-of-study mismatch and overqualification: Labour market correlates and their wage penalty. *IZA Journal of Labor Economics*, 6(2), 1–20. <https://doi.org/10.1186/s40172-016-0052-x>.
- Nguyen, C. M., Koenigs, M., Yamada, T. H., Teo, S. H., Cavanaugh, J. E., Tranel, D., & Denburg, N. L. (2011). Trustworthiness and negative affect predict economic decision making. *Journal of Cognitive Psychology*, 23(6), 748–759. <https://doi.org/10.1080/20445911.2011.575773>
- Palczyńska, M., & Świst, K. (2016). *Measurement properties of non-cognitive scales in the Polish follow-up study on PIAAC (POSTPIAAC)* (OECD Education Working Paper, No. 149).
- Palczyńska, M., & Świst, K. (2018). Personality, cognitive skills and life outcomes: Evidence from the Polish follow-up study to PIAAC. *Large-Scale Assessments in Education*, 6(1), 1–23. <https://doi.org/10.1186/s40536-018-0056-z>.
- Preuss, M., & Hennecke, J. (2018). Biased by success and failure: How unemployment shapes locus of control. *Labour Economics*, 53, 63–74. <https://doi.org/10.1016/j.labeco.2018.05.007>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Rubb, S. (2003). Overeducation in the labor market: A comment and re-analysis of a meta-analysis. *Economics of Education Review*, 22(6), 621–629. [https://doi.org/10.1016/S0272-7757\(02\)00077-8](https://doi.org/10.1016/S0272-7757(02)00077-8)
- Sattinger, M. (1993). Assignment Models of the distribution of earnings. *Journal of Economic Literature*, 31(2), 831–880.
- Sattinger, M. (2012, April). Assignment models and quantitative mismatches. In: *Prepared for the expert workshop "Skill mismatch and firm dynamics: Integrating skills with the world of work"*.
- Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, 51(1), 1–17.
- Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, 98(1), 169–192. <https://doi.org/10.1086/261674>
- Sohn, K. (2010). The role of cognitive and noncognitive skills in overeducation. *Journal of Labor Research*, 31(2), 124–145. <https://doi.org/10.1007/s12122-010-9082-5>
- Strawiński, P., Majchrowska, A., & Broniatowska, P. (2018). Wage returns to different education levels. Evidence from Poland. *Ekonomista*, (1), 25–49.
- Thurow, L. C. (1972). Education and economic equality. *The Public Interest*, 28, 66–81.
- UNESCO Institute for Statistics. (2012). *International standard classification of education: ISCED 2011*. UNESCO Institute for Statistics.
- van der Velden, R., & van Smoorenburg, M. (2000). *The measurement of overeducation and undereducation: self-report vs. job-analyst method*. Les Cahiers du Lasmas, Serie Seminaire (pp. 9–17).
- Verdugo, R. R., & Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *The Journal of Human Resources*, 24(4), 629–643. <https://doi.org/10.2307/145998>
- Verhaest, D., & Omey, E. (2006). The impact of overeducation and its measurement. *Social Indicators Research*, 77(3), 419–448. <https://doi.org/10.1007/s11205-005-4276-6>
- 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(1), 76–90. <https://doi.org/10.1007/s12122-011-9125-6>
- Wincenciak, L. (2016). Educational mismatches and earnings in Poland: Are graduates penalised for being overeducated? . *Ekonomia. Rynek, Gospodarka, Społeczeństwo*, 46, 145–167.

Appendix

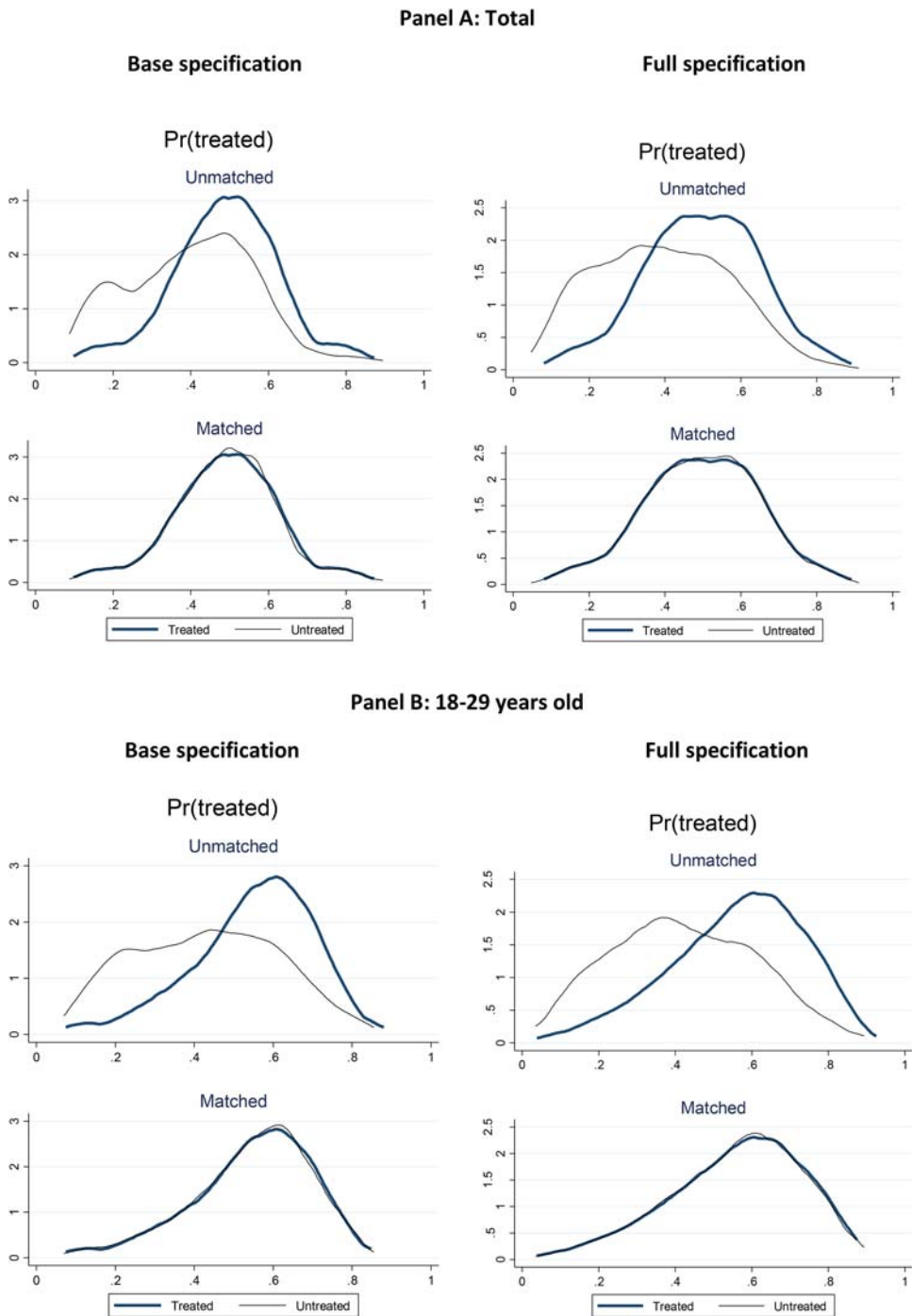
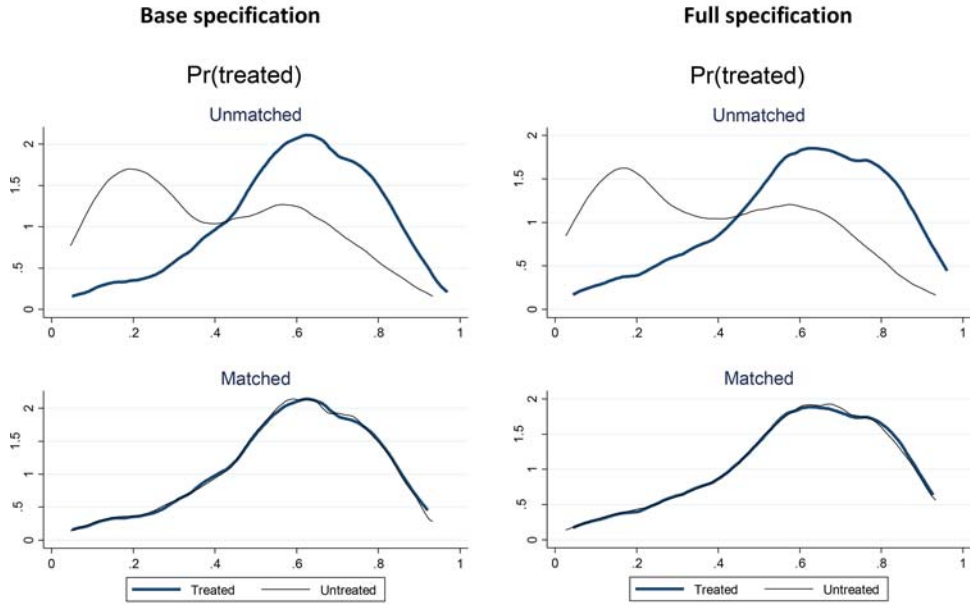


Figure A1. Propensity score distribution. Notes: The graphs present kernel densities (bandwidth=0.06) of the propensity score before and after matching, in the base and the full specification. The base specification includes only socio-demographics, and the full specification includes socio-demographics and cognitive and non-cognitive skills. ^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

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



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

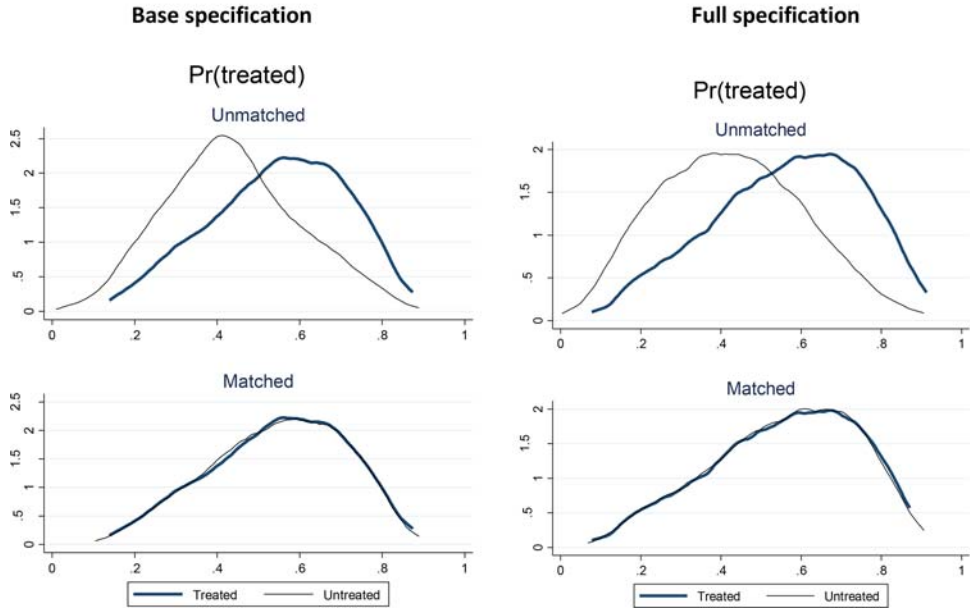


Figure A1 Continued

Panel E: 30-68 years old

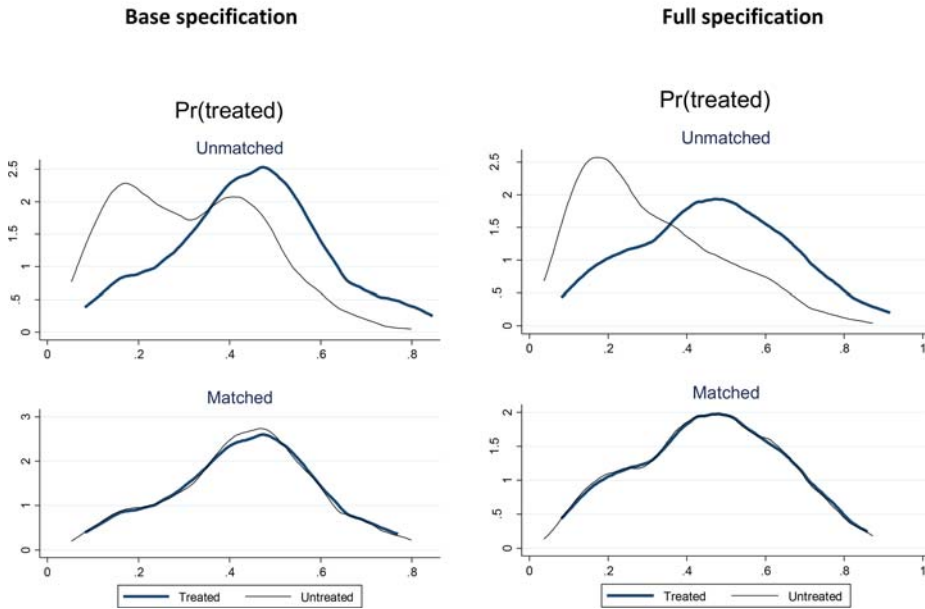


Figure A1 Continued

Table A1. Correspondence table between survey questions on the completed level of education and the required level of education and years of education.

Completed level of education	Required level of education	Completed years of education
Primary (after the reform 1999) – ISCED ^a 1	Primary (after the reform 1999) – ISCED 1	6
Primary (before the reform 1999) or lower secondary school – ISCED 2	Primary (before the reform 1999) or lower secondary school – ISCED 2	8 if year of birth <1986 ^b 9 if year of birth >=1986
Basic vocational (ISCED 3C)	Basic vocational (ISCED 3C)	11
Vocational upper secondary (ISCED 3A)	Vocational upper secondary (ISCED 3A)	13
General upper secondary (ISCED 3A)	General upper secondary (ISCED 3A)	12
Post-secondary non-tertiary (ISCED 4)	Post-secondary non-tertiary (ISCED 4)	14
Bachelor’s degree (ISCED 5A)	Bachelor’s degree (ISCED 5A)	15
Master’s degree or M.D. (ISCED 5A)	Master’s degree or M.D. (ISCED 5A)	17
Doctorate degree (PhD) (ISCED 6)	Doctorate degree (PhD) (ISCED 6)	21

^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

^bCohorts born in 1986 and later were affected by the education system reform in 1999.

Table A2. Differences in mean characteristics between overeducated and matched individuals, full results.

	Mean (overeducated) – Mean (matched)				
	Total	Age 18–29	Age 18–29 (ISCED ^a 3–4)	Age 18–29 (ISCED 5–6)	Age 30–68
Hourly wages (in PLN)	–2.529**	–2.034***	–1.984*	–2.890***	–2.224*
Hourly wages (in PLN) log	–0.144***	–0.133***	–0.134**	–0.180***	–0.123*
Age (in years)	–2.678**	0.059	0.347	–0.236	–1.497
Gender	0.035	0.038	0.099*	–0.045	0.043
<i>Level of education</i>					
ISCED 0–2	–0.015	–0.023*			–0.013
ISCED 3C	–0.164***	–0.139***	–0.288***		–0.160***
ISCED 3A_voc or 4	0.144***	0.125***	0.248***		0.154***
ISCED 3A	0.027	0.023	0.041		0.014
Bachelor's degree	0.02	0.073*		0.144**	–0.018
Master's degree	–0.012	–0.059*		–0.144**	0.023
Private university (for ISCED 5–6)				0.144**	
<i>Field of education</i>					
Teacher training	–0.035	–0.006		–0.015	–0.039
Humanities	0.013	0.03	0.051	0.05	–0.009
Social sciences	0.053	0.045	0.072**	0.002	0.053
Science	0.014	–0.009	0.01	–0.028	0.022
Engineering	–0.002	–0.049	–0.113	0.007	0.028
Agriculture and veterinary	0.006	0.009	0.002	0.01	0.008
Health and welfare	–0.011	0.009	0.052*	–0.037	–0.024
Services	–0.039	–0.029	–0.075	0.012	–0.04
In formal education	–0.009	–0.070*	–0.099	–0.055	–0.01
Work experience (in years)	–3.124***	0.266	0.221	0.414	–2.507*
Living with partner	–0.046	–0.039	0.002	–0.091*	–0.003
Having children aged 0–6	0.015	–0.033	–0.035	–0.022	0.039
<i>Mother's education</i>					
ISCED 0–2	–0.06	0.024	0.045	0.013	–0.061
ISCED 3C	0.001	0.026	0.065	0.017	–0.022
ISCED 3A–4	0.065	0.006	–0.068	0.056	0.08
ISCED 5–6	–0.006	–0.056*	–0.042	–0.086	0.003
<i>Books at home when age 16</i>					
<10	0.032	–0.001	0.009	0.005	0.045
11–100	–0.028	–0.019	–0.04	0.011	–0.03
>100	–0.004	0.021	0.031	–0.016	–0.015
Local unemployment rate	0.43	1.185*	1.282	1.435*	0.036
<i>Cognitive skills</i>					
Literacy	0.028	0.061	0.048	–0.027	–0.005
Numeracy	–0.038	–0.014	–0.07	–0.049	–0.056
ICT (no exp. or failed test)	–0.007	–0.034	–0.036	–0.011	0.013
ICT (refused test)	0.009	0.037	0.043	0.041	0.014
ICT (below level 1)	–0.023	0.001	0.005	0.004	–0.03
ICT (level 1)	–0.049	0.022	0.02	0.011	–0.085*
ICT (level 2 and 3)	0.070*	–0.026	–0.033	–0.044	0.087
<i>Personality traits</i>					
B5: Conscientiousness	–0.086	–0.08	0.01	–0.127*	–0.044
B5: Extraversion	–0.149*	–0.1	–0.06	–0.101	–0.175
B5: Agreeableness	–0.106	–0.01	0.045	–0.033	–0.106
B5: Openness	–0.160*	–0.091	–0.091	–0.056	–0.17
B5: Neuroticism	0.055	–0.036	0.04	–0.115	0.066
Grit	–0.068	–0.017	0.1	–0.15	–0.068
Observations	1598	1091	512	548	507

Notes: ***/**/* indicate statistical significance at the 0.1%/1%/5%-level. Literacy, numeracy, and non-cognitive skills are standardized.

^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

Table A3. Selection into overeducation (logit coefficients), full results.

	Total	Age 18–29	Age 18–29 (ISCED ^a 3–4)	Age 18–29 (ISCED 5–6)	Age 30–68
Age	0.083 (0.081)	–0.416 (0.636)	–1.762* (0.882)	2.738 (1.610)	0.317 (0.168)
Age ^2	–0.001 (0.001)	0.006 (0.013)	0.034 (0.018)	–0.057 (0.031)	–0.003 (0.002)
Female	–0.027 (0.209)	0.039 (0.154)	–0.149 (0.270)	0.135 (0.250)	–0.080 (0.309)
<i>Level of education</i>					
ISCED 0–2	–2.130** (0.643)	–2.003** (0.728)			–2.302* (0.995)
ISCED 3C	–1.743** (0.568)	–1.740** (0.541)	–3.194** (0.979)		–1.874* (0.871)
ISCED 3A_voc or 4	0.147 (0.465)	0.753* (0.367)	–0.372 (0.808)		0.017 (0.782)
ISCED 3A	ref.	ref.	ref.		ref.
Bachelor's degree	–0.122 (0.451)	0.859* (0.353)		0.444 (0.271)	–0.884 (0.921)
Master's degree	–0.255 (0.454)	0.301 (0.369)		ref.	–0.303 (0.758)
Informal education	–0.508 (0.267)	–0.775*** (0.213)	–0.787* (0.392)	–0.619* (0.280)	–0.416 (0.734)
<i>Field of education</i>					
Teacher training	–0.309 (0.371)	–0.502 (0.416)		–0.647 (0.487)	–0.140 (0.504)
Humanities	ref.	ref.	ref.	ref.	ref.
Social sciences	0.261 (0.376)	–0.272 (0.331)	1.628 (1.065)	–0.757 (0.410)	0.457 (0.553)
Science, mathematics, and computer science	0.173 (0.421)	–0.561 (0.388)	0.676 (0.866)	–0.907 (0.478)	0.570 (0.623)
Engineering	0.312 (0.389)	–0.383 (0.337)	0.497 (0.801)	–0.344 (0.443)	0.599 (0.582)
Agriculture and veterinary	0.159 (0.634)	–0.133 (0.482)	0.371 (1.017)	0.269 (0.636)	0.387 (0.853)
Health and welfare	–0.428 (0.383)	–0.461 (0.379)	1.078 (1.029)	–1.025* (0.508)	–0.527 (0.872)
Services	0.057 (0.431)	–0.505 (0.444)	0.536 (0.923)	–0.392 (0.650)	0.276 (0.618)
Work experience (in years)	–0.078 (0.042)	0.147** (0.052)	0.146 (0.089)	0.275** (0.099)	–0.152* (0.059)
Work experience ^2	0.001 (0.001)	–0.004* (0.002)	–0.003 (0.002)	–0.012 (0.006)	0.003* (0.001)
Living with partner = 1	–0.086 (0.215)	–0.298 (0.166)	–0.174 (0.294)	–0.422 (0.243)	–0.049 (0.321)
Having children aged 0–6 = 1	–0.054 (0.233)	–0.039 (0.222)	–0.324 (0.417)	0.194 (0.326)	0.038 (0.327)
<i>Mother's education</i>					
ISCED 0–2	ref.	ref.	ref.	ref.	ref.
ISCED 3C	0.071 (0.262)	–0.381 (0.278)	–0.270 (0.447)	–0.181 (0.574)	0.113 (0.330)
ISCED 3A-4	0.242 (0.258)	–0.554 (0.294)	–0.855 (0.471)	0.097 (0.547)	0.476 (0.300)
ISCED 5–6	–0.048 (0.349)	–0.904* (0.386)	–1.450 (0.800)	–0.303 (0.576)	0.151 (0.508)
<i>Books at home when age 16</i>					
<10	ref.	ref.	ref.	ref.	ref.
11–100	–1.003** (0.373)	–0.330 (0.420)	–0.273 (0.512)	–0.644 (1.286)	–1.146* (0.497)
>100	–0.976* (0.432)	0.011 (0.451)	0.171 (0.589)	–0.326 (1.246)	–1.234* (0.603)
Local unemployment rate	0.016 (0.021)	0.045** (0.016)	0.055 (0.032)	0.049* (0.019)	0.001 (0.032)

(Continued)

Table A3. Continued.

	Total	Age 18–29	Age 18–29 (ISCED ^a 3–4)	Age 18–29 (ISCED 5–6)	Age 30–68
Numeracy	–0.234* (0.108)	–0.042 (0.104)	–0.169 (0.162)	0.052 (0.153)	–0.263 (0.153)
ICT (no exp. or failed test)	ref.	ref.	ref.	ref.	ref.
ICT (refused test)	0.119 (0.399)	1.059* (0.465)	1.308* (0.656)	0.852 (0.688)	0.006 (0.466)
ICT (below level 1)	–0.433 (0.424)	0.569 (0.403)	0.478 (0.570)	0.664 (0.865)	–0.551 (0.535)
ICT (level 1)	–0.582 (0.346)	0.481 (0.345)	0.438 (0.444)	0.392 (0.713)	–0.795 (0.455)
ICT (levels 2 and 3)	0.028 (0.367)	0.422 (0.336)	0.207 (0.422)	0.423 (0.649)	0.185 (0.483)
B5: Conscientiousness	0.055 (0.222)	–0.458* (0.209)	–0.105 (0.377)	–0.697* (0.286)	0.322 (0.336)
B5: Extraversion	–0.122 (0.143)	–0.163 (0.145)	–0.057 (0.233)	–0.142 (0.168)	–0.140 (0.178)
B5: Agreeableness	0.025 (0.215)	0.577* (0.236)	0.208 (0.397)	0.878** (0.313)	–0.198 (0.326)
B5: Openness	–0.107 (0.166)	–0.143 (0.158)	–0.269 (0.256)	–0.179 (0.190)	–0.151 (0.223)
B5: Neuroticism	0.095 (0.105)	–0.002 (0.080)	0.117 (0.131)	–0.105 (0.138)	0.135 (0.148)
Grit	–0.057 (0.100)	–0.084 (0.086)	–0.067 (0.134)	–0.081 (0.122)	–0.021 (0.138)
Private university				0.616* (0.277)	
Constant	–0.099 (1.458)	5.801 (8.137)	21.867* (10.832)	–33.526 (20.989)	–4.605 (3.792)
Observations	1598	1091	512	548	507

Notes: Logit coefficients. Standard errors are in parentheses.

***/**/*Indicate statistical significance at the 0.1%/1%/5%-level. Numeracy and non-cognitive skills are standardized. ^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

Table A4. Matching quality.

Total	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	12.08	7.78	0.08	0.04	10.14	6.68	0.09	0.04
NN_4	12.08	6.49	0.08	0.02	10.14	5.02	0.09	0.03
Kernel_0.02	12.08	5.29	0.08	0.01	10.14	5.32	0.09	0.02
Kernel_0.06	12.08	5.65	0.08	0.02	10.14	5.42	0.09	0.02
Kernel_0.2	12.08	6.77	0.08	0.02	10.14	6.00	0.09	0.03
Age 18–29	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	10.57	6.10	0.09	0.02	8.78	5.31	0.10	0.02
NN_4	10.57	3.12	0.09	0.01	8.78	4.66	0.10	0.02
Kernel_0.02	10.57	2.76	0.09	0.01	8.78	3.93	0.10	0.01
Kernel_0.06	10.57	2.71	0.09	0.00	8.78	3.84	0.10	0.01
Kernel_0.2	10.57	3.47	0.09	0.01	8.78	3.22	0.10	0.01
Age 18–29 (ISCED ^a 3–4)	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	13.51	11.03	0.14	0.06	10.11	12.02	0.15	0.09
NN_4	13.51	7.62	0.14	0.02	10.11	7.76	0.15	0.04
Kernel_0.02	13.51	6.78	0.14	0.02	10.11	7.96	0.15	0.04
Kernel_0.06	13.51	5.42	0.14	0.01	10.11	7.03	0.15	0.03
Kernel_0.2	13.51	5.17	0.14	0.02	10.11	6.00	0.15	0.03
Age 18–29 (ISCED 5–6)	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	11.12	9.03	0.06	0.04	10.38	6.92	0.08	0.05
NN_4	11.12	5.08	0.06	0.01	10.38	5.83	0.08	0.02
Kernel_0.02	11.12	4.01	0.06	0.01	10.38	5.47	0.08	0.01
Kernel_0.06	11.12	3.46	0.06	0.01	10.38	4.87	0.08	0.01
Kernel_0.2	11.12	4.58	0.06	0.01	10.38	4.45	0.08	0.01
Age 30–68	Base				Full			
	MB_bef	MB_aft	r2bef	r2aft	MB_bef	MB_aft	r2bef	r2aft
NN_1	11.01	7.94	0.10	0.04	11.06	8.21	0.12	0.08
NN_4	11.01	4.27	0.10	0.01	11.06	5.48	0.12	0.03
Kernel_0.02	11.01	4.34	0.10	0.01	11.06	5.22	0.12	0.02
Kernel_0.06	11.01	3.26	0.10	0.01	11.06	4.13	0.12	0.01
Kernel_0.2	11.01	4.52	0.10	0.01	11.06	4.59	0.12	0.02

Notes: Matching algorithms: NN_1: nearest neighbour (NN) matching with 1 neighbour; NN_4: NN matching with 4 neighbours; Kernel_0.02: Epanechnikov kernel matching with a bandwidth of 0.02; Kernel_0.06: Epanechnikov kernel matching with a bandwidth of 0.06; Kernel_0.2: Epanechnikov kernel matching with a bandwidth of 0.2. Quality measures: MB_bef: the mean absolute standardised bias before matching; MB_aft: the mean absolute standardised bias after matching; r2bef: Pseudo R2 from probit estimation of the propensity score on all the variables on raw samples; r2aft: Pseudo R2 from probit estimation of the propensity score on all the variables on matched samples.

^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).



Table A5. OLS results.

	Total		Age 18–29		Age 18–29 (ISCED ^a 3–4)		Age 18–29 (ISCED 5–6)		Age 30–68	
	Base	Full	Base	Full	Base	Full	Base	Full	Base	Full
Overeducated	-0.141*** (0.033)	-0.134*** (0.032)	-0.132*** (0.032)	-0.128*** (0.032)	-0.151** (0.054)	-0.133* (0.054)	-0.155*** (0.035)	-0.157*** (0.035)	-0.137** (0.042)	-0.135** (0.042)
Numeracy	0.034 (0.017)	0.034 (0.017)	0.052** (0.020)	0.052** (0.020)	0.038 (0.024)	0.038 (0.024)	0.044 (0.025)	0.044 (0.025)	0.027 (0.021)	0.027 (0.021)
B5: Conscientiousness	0.095* (0.037)	0.095* (0.037)	0.023 (0.033)	0.023 (0.033)	-0.019 (0.050)	-0.019 (0.050)	0.086 (0.049)	0.086 (0.049)	0.119* (0.050)	0.119* (0.050)
B5: Extraversion	0.026 (0.021)	0.026 (0.021)	0.024 (0.021)	0.024 (0.021)	0.037 (0.035)	0.037 (0.035)	0.022 (0.027)	0.022 (0.027)	0.039 (0.026)	0.039 (0.026)
B5: Agreeableness	-0.107* (0.042)	-0.107* (0.042)	-0.029 (0.039)	-0.029 (0.039)	0.005 (0.046)	0.005 (0.046)	-0.067 (0.057)	-0.067 (0.057)	-0.129* (0.054)	-0.129* (0.054)
B5: Openness	-0.013 (0.026)	-0.013 (0.026)	-0.008 (0.030)	-0.008 (0.030)	-0.014 (0.040)	-0.014 (0.040)	0.000 (0.036)	0.000 (0.036)	-0.022 (0.033)	-0.022 (0.033)
B5: Neuroticism	-0.049** (0.014)	-0.049** (0.014)	-0.035* (0.017)	-0.035* (0.017)	-0.016 (0.023)	-0.016 (0.023)	-0.048* (0.022)	-0.048* (0.022)	-0.062** (0.019)	-0.062** (0.019)
Grit	-0.019 (0.014)	-0.019 (0.014)	-0.001 (0.014)	-0.001 (0.014)	0.012 (0.016)	0.012 (0.016)	-0.016 (0.020)	-0.016 (0.020)	-0.020 (0.018)	-0.020 (0.018)
Observations	1598	1598	1091	1091	512	512	548	548	507	507
R ²	0.420	0.440	0.208	0.228	0.149	0.177	0.205	0.235	0.477	0.503

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 for still in education, ICT levels, experience, experience squared, a dummy for living with a partner, a dummy for having children aged 0–6, mother's level of education and number of books in the home when the respondent was age 16, local unemployment rate; in the 18–29 (ISCED 5–6) subpopulation, additionally a dummy for attending a private university.

^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).

Table A6. OLS results: specification with an undereducation dummy.

	Total		Age 18-29		Age 18-29 (ISCED ^a 3-4)		Age 18-29 (ISCED 5-6)		Age 30-68	
	base	full	base	full	base	full	base	full	base	full
Overeducated	-0.143*** (0.031)	-0.134*** (0.031)	-0.128*** (0.032)	-0.123*** (0.032)	-0.151*** (0.054)	-0.137* (0.054)	-0.149*** (0.036)	-0.148*** (0.037)	-0.143*** (0.039)	-0.135*** (0.039)
Undereducated	0.038 (0.037)	0.041 (0.036)	-0.009 (0.050)	0.004 (0.049)	-0.039 (0.059)	-0.021 (0.055)	0.168 (0.139)	0.176 (0.160)	0.052 (0.049)	0.055 (0.048)
Numeracy		0.039* (0.017)		0.048* (0.018)		0.032 (0.024)		0.039 (0.026)		0.038 (0.021)
B5: Conscientiousness		0.094*** (0.033)		0.030 (0.032)		0.028 (0.050)		0.092 (0.049)		0.114* (0.044)
B5: Extraversion		0.020 (0.018)		0.019 (0.019)		0.024 (0.027)		0.026 (0.028)		0.029 (0.023)
B5: Agreeableness		-0.113** (0.038)		-0.045 (0.036)		-0.048 (0.046)		-0.074 (0.057)		-0.133** (0.048)
B5: Openness		-0.006 (0.023)		-0.001 (0.027)		-0.008 (0.031)		-0.001 (0.035)		-0.012 (0.028)
B5: Neuroticism		-0.043** (0.013)		-0.027 (0.016)		-0.012 (0.020)		-0.043 (0.023)		-0.055** (0.017)
Grit		-0.013 (0.013)		0.006 (0.013)		0.014 (0.014)		-0.017 (0.020)		-0.015 (0.017)
Observations	1818	1818	1220	1220	608	608	558	558	598	598
R ²	0.400	0.421	0.204	0.219	0.146	0.165	0.208	0.233	0.454	0.480

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 for still in education, ICT levels, experience, experience squared, a dummy for living with a partner, a dummy for having children aged 0-6, mother's level of education and number of books in the home when the respondent was age 16, local unemployment rate; in the 18-29 (ISCED 5-6) subpopulation, additionally a dummy for attending private university.

^aInternational Standard Classification of Education 1997 (UNESCO Institute for Statistics, 2012).