Social Indicators Research (2025) 177:763–786
https://doi.org/10.1007/s11205-025-03540-x

ORIGINAL RESEARCH

Gender-Specific Analysis of Self-Reported Health and Educational Mismatch: Evidence from Employees in Russia

Silvana Robone[1,2,3,4]

Accepted: 27 January 2025 / Published online: 8 February 2025

Abstract

Our paper investigates the influence of vertical educational mismatch—overeducation and undereducation—on selected EQ-5D metrics, namely pain and anxiety/depression. We conduct a gender-specific analysis and estimate ordered probit models for our categorical dependent variables by using a sample of currently working employees from the Russia Longitudinal Monitoring Survey. Since our health outcomes are self-reported, we challenge the validity of the results obtained by using the ordered probit model and enrich the analysis by correcting our estimates for the presence of reporting heterogeneity bias. To do that, we merge the RLMS-HSE (2005) with externally collected anchoring vignettes for Russia from the World Health Survey (2003) and estimate a hierarchical ordered probit (HOPIT) model. The HOPIT estimates show that in several cases, educational mismatch affects the reporting style of respondents. Our findings provide evidence that, after adjusting for reporting heterogeneity, undereducation has a negative influence on the physical component of health (proxied by pain) for women, while overeducation affects the psychological component of health (proxied by anxiety/depression) in both gender groups. Overeducated women appear to have better psychological health, while overeducated men have worse psychological health than their matched counterparts.

Keywords Educational mismatch · Overeducation · Undereducation · EQ-5D · RLMS-HSE · World health survey · Reporting heterogeneity bias · Anchoring vignettes · HOPIT model

JEL Classification I10 · I26 · I29 · J01 · J24 · J80 · J81

✉ Silvana Robone
silvana.robone@uniupo.it

[1] Department of Law and Political, Economic and Social Sciences, University of Eastern Piedmont, Via Cavour 84, 15121 Alessandria, Italy

[2] Econometrics and Data Group, HEDG - Health, University of York, York, UK

[3] Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University, Milan, Italy

[4] Department of Economics and Finance, Università Cattolica del Sacro Cuore Milano, Milan, Italy

1 Introduction

Over the past decades the policy-making institutions in low- and middle-income countries have persistently considered a well-educated working population as a precondition for sustainable economic growth and development (Handel et al., 2016). This idea comes from the human capital theory (Becker, 1993; Schultz, 1963) which implies that individuals accumulate a stock of knowledge via learning-or-doing (i.e. schooling) and learning-by-doing processes which contribute to an increase in their productivity and, consequently, determine the long-term economic growth of their country (Mankiw et al., 1992; Romer, 1990). However, recent studies, which consider education from an economic perspective, argue that sustainable growth is rather associated with an appropriate distribution of employees over occupations with respect to the level of knowledge and skills they have, than with their achieved educational attainment (Handel et al., 2016).

The comparison of outcomes of the national systems of professional and vocational education (represented by the actual level of knowledge and skills employees obtain through education) with the labour market needs (represented by employment requirements related to the level of knowledge and skills of potential employees) indicates the presence of structural inefficiencies also known as educational mismatch. Being caused by either labour market imbalances (e.g. excess labour demand, risk and uncertainty etc.) or institutional factors (e.g. technological change, short-term recessions etc.), the incidence of educational mismatch—divergence of the level of education/skills the employee actually has from the one required for the position he/she holds—is observed in both developed and developing countries and tends to remain persistent over time (Davia et al., 2017; Ghaffarzadegan et al., 2017; Morgado et al., 2016). For instance, according to the Report of the European Centre for the Development of Vocational Training (2018), about 29 per cent of EU adults experience various forms of educational mismatch in the labour market. Moreover, the number of mismatched employees in the EU economy is expected to increase during the next few years due to a slowdown in job growth, a high unemployment rate and expansion of low-quality jobs accompanied by the prevalence of tertiary education among the European population (Cedefop, 2018).

The previous studies on educational mismatch mainly consider this structural inefficiency in the context of labour market outcomes, e.g. wages, returns to education, productivity of firms (Iriondo & Perez-Amaral, 2016; Kampelmann & Rycx, 2012; Pecoraro, 2016; Romero et al., 2017; Sellami et al., 2017; Verhaest & Omey, 2012), occupational mobility and career dynamics (Acosta-Ballesteros et al., 2017; Baert et al., 2013; Diem & Wolter, 2014; Kiersztyn, 2013; Mavromaras & McGuinness, 2012; Meroni & Vera-Toscano, 2017; Vossemer & Schuck, 2015). There are also a few studies which investigate the association between educational mismatch and psychological well-being and health-related outcomes of the labour force. They are mainly dedicated to the impact of overeducation on job and life satisfaction (Artes et al., 2013; Badillo-Amador & Vila, 2013; Mateos-Romero & Salinas-Jimenez, 2018; Piper, 2015; Salinas-Jimenez et al., 2016; Ueno & Krause, 2018) and mental health (Bracke et al., 2014; Hardie, 2014; Milner et al., 2017; Mossakowski, 2011; Zhu & Chen, 2016) and alcohol consumption (Robone, 2023).

This study seeks to contribute to the literature which considers educational mismatch as a determinant of mental health and psychological well-being of employees. In our analysis, we investigate the influence of both types of vertical educational mismatch—overeducation and undereducation—on the physical and psychological components of general health, as proxied by pain and anxiety/depression, respectively. We conduct a gender-specific analysis and estimate ordered probit (OPROBIT) models for the selected EQ-5D metrics on the sample of currently working employees from the Russia Longitudinal Monitoring Survey (RLMS-HSE). Since our health outcomes are self-reported, we challenge the validity of our preliminary results and enrich the analysis by testing for the presence of reporting heterogeneity bias. To do that, we merge the RLMS-HSE (2005) with externally collected anchoring vignettes for Russia from the World Health Survey (2003) and estimate the hierarchical ordered probit (HOPIT) models, as suggested in the methodology proposed by Harris et al. (2020).

Our results provide evidence that educational mismatch differently affects the health of men and women in Russia. More precisely, undereducation tends to be associated with pain in the female sub-sample, while overeducation is related to anxiety/depression in the male employees. In addition, both types of vertical educational mismatch have an impact on the way employees report their health status (reporting styles). Overall, our findings reveal that vertical educational mismatch affects both the psychological (proxied by anxiety/depression) and the physical (proxied by pain) component of general health.

Our paper contributes to the current state of socioeconomic research on educational mismatch in several ways. First, we fill the gap in the literature by estimating the influence of *both* types of vertical educational mismatch—overeducation and undereducation—with the health outcomes of employees. Moreover, we focus our analysis on Russia. This is one of the countries where educational mismatch raises serious concerns since "large imbalances between the supply and demand for skills… are driven by rapid economic restructuring, misalignment of the education system with labor market needs, and underdeveloped adult education and training systems" (Kupets, 2016b, p.1). Finally, to the best of our knowledge, this is the first study which considers the issue of reporting heterogeneity in the context of educational mismatch.

2 Related Literature

The recent literature on educational mismatch is mainly focused on the incidence, determinants and socioeconomic outcomes of overeducation in the labour market. These studies provide evidence that employees who experience this type of mismatch are more likely to face certain constraints in terms of employment in comparison with their matched counterparts. More precisely, overeducated employees seem to be disadvantaged in terms of remuneration and career advancement which may cause a substantial decline in perceived levels of well-being and health. For instance, Ueno and Krause (2018) and Piper (2015) provide evidence of the negative impact of overeducation on job and life satisfaction, respectively. The scholars conclude that individuals consider costs associated with education as a sort of investment in more promising employment opportunities in the future. Consequently, if their expectations are not properly met, they tend to suffer from emotional distress leading to a substantial decline in job satisfaction and overall well-being.

By using data from the European Social Survey (ESS), Salinas-Jimenez et al. (2016) and Artes et al. (2013) demonstrate that the number of years of overeducation (in Salinas-Jimenez et al., 2016) and overeducation itself (in Artes et al., 2013) are associated with a significant decline in life satisfaction while the opposite can be observed for undereducation. Like the previous study, by using the data from the European Social Survey (ESS), Bracke et al. (2014) shows that although education itself creates certain benefits for the mental health of employees, these benefits tend to drop drastically or even disappear in the cases when employees have to deal with a surplus of human capital—overeducation.

The study of Hultin et al. (2016) uses data from the Stockholm Public Health Cohort (2010) and estimates the effect of overeducation on self-assessed health (SAH) and psychological distress. The researchers reveal the significant negative impact of overeducation on SAH in the female sub-sample, while no effect is observed among the male employees. Finally, Garcy (2015) investigates the impact of prolonged vertical educational mismatch on mortality in Sweden and reports that overeducated employees are more prone to excessive mortality, while undereducation tends to provide a protective effect against mortality among native-born Swedish employees of both gender groups.

Overall, the limited number of studies on health-related outcomes of educational mismatch (mainly overeducation) agree on the point that substantial decline in the levels of well-being and health of mismatched employees can be explained from the perspective of improperly met career expectations. Of the recently published papers, most tackle the issue of educational mismatch in the context of developed economies, almost completely overlooking the group of low- and middle-income countries (LMICs) and late-reforming transition economies (both included in the EU and non-EU members). However, the latter countries could be particularly interesting for the research on educational mismatch. They experienced a prolonged period with a communist regime in the past and, as a result, are characterized by large imbalances between the supply and demand for skills, mainly driven by the misalignment of educational systems with the needs of the labour market and poorly performing systems of adult and vocational training needed nowadays (Kupets, 2016a; Kyui, 2010). In particular, among OECD countries, the highest levels of decline in job satisfaction related to both types of vertical educational mismatch are detected in Eastern Europe and Asia (Mateos-Romero & Salinas-Jimenez, 2018).

3 Research Methodology

3.1 Descriptive Analysis: Data, Variables and Sample

In order to investigate the association between educational mismatch and selected health outcomes, we use individual data from the 14th wave of the Russia Longitudinal Monitoring Survey (RLMS-HSE). The questionnaires of the RLMS-HSE are designed to be compatible with the well-established European national surveys such as the British Household Panel Survey (BHPS) and the German Socio-Economic Panel (GSOEP). They contain a wide range of data on health, well-being, educational attainment, employment status and welfare of the Russian population, collected at both the individual and the household level.

Our principal sample consists of 3,390 Russian respondents, who are currently employed and are in the working-age population.[1] We exclude from the analysis retired, unemployed and disabled respondents as well as those who work unofficially, have a second job or are on maternity leave at the time of the survey.[2] Armed forces personnel and

[1] We adjust for the existing gender differences in retirement age in Russia in 2005 by cutting our sample at 54 and 59 years for women and men, respectively.

[2] Since a period of maternity leave may last up to 3 years in Russia (first half with a compensation from the employer and the last one with a government subsidy), we exclude these respondents from the analysis due to their weak participation in working activities over this period.

entrepreneurs are also removed from the sample due to the very specific tasks and relatively small number of observations for these professional groups.

Our self-reported dependent variables are represented by the selected EQ-5D metrics of pain and anxiety/depression.[3] They are defined in the RLMS-HSE by the questions 'Do you feel any pain?' and 'Do you feel any anxiety/depression?' and measured on a three-point categorical scale. In order to simplify the interpretation of our estimates, we re-scale the original variables from the worst to the best health outcome in such a way that the first category represents "acute", the second "some", and the third "none" for each self-reported status.

Our main explanatory variables of overeducation and undereducation, are dummy variables constructed with respect to the method of realized matches, originally introduced by Verdugo and Verdugo (1989) and implemented in a wide range of studies (McGuinness et al., 2017). We start by defining the mean value of years of completed education computed by employees in each occupation-industry group and then set an interval from -1 standard deviation to $+1$ standard deviation around the computed mean.[4] For the next step, we compare the actual number of years of completed education the respondent has with the values from the earlier specified interval. Those respondents whose actual number of years of completed education exceeds (is less than) the upper (lower) boundary of the interval are classified as overeducated (undereducated). If the respondent's actual number of years of education belongs to the specified interval, he/she is defined as vertically matched in terms of education.

We also use a set of standard sociodemographic covariates in our models, as done in the previous studies (Marmot et al., 2012). We include age (de-meaned and divided by 10) and its squared term to test for the U-shaped relationship. Marital status is defined as a dummy variable where married/co-habiting respondents are taken as a reference category, while never married, divorced/separated individuals and widowers are combined under the category "single". Educational attainment is represented by a dummy variable which equals 1 if the respondent has a completed tertiary education and 0 otherwise.[5] In order to take into account the economic status of the respondents, we use permanent income indicators from the household part of the RLMS-HSE (2005). We construct income quintiles by means of principal component analysis and use the third quintile (medium income group) as a reference category in our analysis (Ferguson et al., 2003). In addition, we control for frequent

[3] EQ-5D is a generic instrument for measuring health status which covers five dimensions—mobility, self-care, usual activities, pain/discomfort and anxiety/depression. It has been implemented in many population health surveys and widely used in the clinical studies and economic evaluation of health care (Herdman et al., 2011; Rabin and de Charro, 2001). We do not use mobility, self-care and usual activities as dependent variables in our analysis since the respondents who experience difficulties in these dimensions are more likely to be assigned to a disability class and, as a result, to be restricted in terms of labour market participation.

[4] Occupational categories are taken from the International Standard Classification of Occupations (ISCO-08) and include Managers, Professionals, Technicians and Associate Professionals, Clerical Support Workers, Services and Sales Workers, Skilled Agricultural, Forestry and Fishery Workers, Craft and Related Trades Workers, Plant and Machine Operators and Assemblers, and Elementary Occupations. Industries are represented by Manufacturing, Agriculture and Forestry, Construction, Transportation and Communication, Trade, Public Administration (including Army and Church), Business Services, and Social Services.

[5] The lack of control for education in the model leads to the situation when its effect is captured by educational mismatch. By using a dummy variable which arises from categories of educational attainment, we keep the correlation coefficients between this control and our mismatch variables in the range of 0.2–0.3, thus avoiding the problem of multicollinearity in our models.

visits to a GP in the past 12 months. The respondents who visited a doctor less often than 2–3 times a year are considered as a reference category for this dummy variable. Finally, we include a control for living in a metropolitan area which equals 1 if the respondent is settled either in Moscow or St. Petersburg.

Along with sociodemographic covariates, we include some work-related controls in our models, since we believe that they may affect both the health status of respondents and their probability of ending up as mismatched employees in the labour market. More precisely, we include controls for the company's size and its type of ownership, along with the respondent's employment history over the past 12 months. We specify the company's size as a dummy variable where small and medium enterprises (with less than 250 employees) are taken as a reference category. The type of company ownership is also a dummy variable which equals 1 if the company is (at least partially) owned by the government. The employment history of each respondent is represented by a set of dummies for occupational change and unemployment (experienced in the last 12 months) which equal 1 if the described situations have occurred and 0 otherwise. In this case, the respondents who have not changed their employment status in the past year serve as the reference category.

Table 1 provides descriptive statistics for the independent variables stratified by gender. As the data depict, the most striking gender differences are observed for such variables as marital status, educational attainment, frequent visits to a GP in the last 12 months, company's type of ownership and occupational change experienced in the past 12 months. The share of respondents whose level of education exceeds the one

Table 1 Descriptive statistics for independent variables stratified by gender: t-test of equal means

Variable	Women (n = 1,732)		Men (n = 1,658)		*p* value
	Mean	Std. Dev	Mean	Std. Dev	
Sociodemographic controls:					
Age	37.260	9.662	37.000	10.799	
Single	0.318	0.466	0.201	0.401	***
Education	0.656	0.475	0.415	0.493	***
Low income group	0.201	0.401	0.201	0.401	
Medium low income group	0.199	0.400	0.203	0.402	
Medium high income group	0.202	0.401	0.207	0.406	
High income group	0.193	0.395	0.190	0.392	
Frequent visits to GP (12 m)	0.463	0.499	0.233	0.423	***
Metropolitan area	0.126	0.332	0.128	0.335	
Work-related controls:					
Large company	0.210	0.408	0.246	0.431	
Public sector	0.605	0.489	0.507	0.500	***
Occupational change (12 m)	0.137	0.344	0.213	0.409	***
Unemployment (12 m)	0.049	0.216	0.036	0.187	*
Main explanatory variables:					
Overeducation	0.133	0.339	0.156	0.363	*
Undereducation	0.171	0.377	0.156	0.363	

*$p<0.10$, **$p<0.05$, ***$p<0.01$

required for their position is quite comparable for both gender groups and equals 13.3 and 15.6 per cent for women and men, respectively. Finally, the share of undereducated women equals 17.1 per cent while undereducated men account only for 15.6 per cent.

Tables 2 and 3 report descriptive statistics for the women and men sub-samples, respectively, which are stratified by categories of educational mismatch. Overeducated respondents differ from their matched counterparts in terms of educational attainment and income (both gender groups) and employment history (in the female sub-sample). Undereducated and matched women reveal statistically significant differences in terms of age, educational attainment, income and company characteristics. As for men, undereducated and matched respondents differ in terms of educational attainment, income, and company size.

Figures 1 and 2 display how self-reports on pain and anxiety/depression are distributed over categories of educational mismatch and gender. At first glance, mismatched women tend to report the "acute" category of pain more frequently than their matched counterparts. In contrast, we do not observe any significant differences in pain reporting between mismatched and matched male respondents. Furthermore, undereducated women and overeducated men do not seem to differ significantly in terms of self-reports on anxiety/depression from their matched counterparts. In contrast, overeducated women choose "acute" and "none" categories of anxiety/depression more frequently than matched respondents of the same gender, while the opposite holds for the "some" category. Finally, undereducated men tend to choose the

Table 2 Descriptive statistics stratified by categories of educational mismatch: female sub-sample

	Matched (n = 1,205)		Overeducated (n = 230)			Undereducated (n = 297)		
	Mean	Std. Dev	Mean	Std. Dev	*p* value	Mean	Std. Dev	*p* value
Sociodemographic controls:								
Age	37.101	9.684	36.091	9.387		38.811	9.630	***
Single	0.319	0.466	0.313	0.465		0.316	0.466	
Education	0.659	0.474	0.943	0.231	***	0.424	0.495	***
Low income group	0.194	0.396	0.170	0.376		0.253	0.435	**
Medium low income group	0.203	0.403	0.191	0.394		0.189	0.392	
Medium high income group	0.209	0.407	0.209	0.407		0.165	0.372	*
High income group	0.183	0.386	0.270	0.445	***	0.175	0.381	
Frequent visits to GP (12 m)	0.475	0.500	0.426	0.496		0.444	0.498	
Metropolitan area	0.120	0.325	0.148	0.356		0.138	0.346	
Work-related controls:								
Large company	0.221	0.415	0.200	0.401		0.175	0.381	*
Public sector	0.592	0.492	0.617	0.487		0.646	0.479	*
Occupational change (12 m)	0.131	0.338	0.178	0.384	*	0.128	0.335	
Unemployment (12 m)	0.047	0.212	0.083	0.276	**	0.030	0.172	

Overeducated and undereducated female respondents are compared with matched ones (t-test of equal means)

*$p<0.10$, **$p<0.05$, ***$p<0.01$

Table 3 Descriptive statistics stratified by categories of educational mismatch: male sub-sample

	Matched (n = 1,141)		Overeducated (n = 259)			Undereducated (n = 258)		
	Mean	Std. Dev	Mean	Std. Dev	*p* value	Mean	Std. Dev	*p* value
Sociodemographic controls:								
Age	36.854	10.581	37.097	10.880		37.550	11.662	
Single	0.193	0.395	0.228	0.420		0.213	0.410	
Education	0.387	0.487	0.807	0.395	***	0.143	0.351	***
Low income group	0.192	0.394	0.178	0.383		0.267	0.443	***
Medium low income group	0.213	0.410	0.178	0.383		0.182	0.387	
Medium high income group	0.208	0.406	0.224	0.418		0.190	0.393	
High income group	0.184	0.388	0.263	0.441	***	0.143	0.351	
Frequent visits to GP (12 m)	0.231	0.422	0.274	0.447		0.202	0.402	
Metropolitan area	0.128	0.334	0.131	0.338		0.128	0.335	
Work-related controls:								
Large company	0.255	0.436	0.278	0.449		0.174	0.380	***
Public sector	0.503	0.500	0.525	0.500		0.508	0.501	
Occupational change (12 m)	0.204	0.403	0.224	0.418		0.240	0.428	
Unemployment (12 m)	0.039	0.195	0.035	0.183		0.023	0.151	

Overeducated and undereducated male respondents are compared with matched ones (t-test of equal means)
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

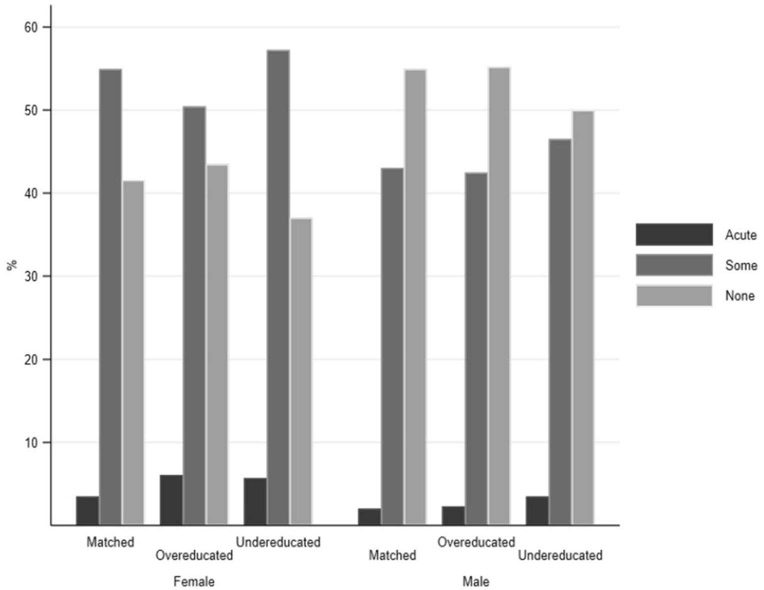

Fig. 1 Distribution of self-reports on pain over categories of educational mismatch and gender

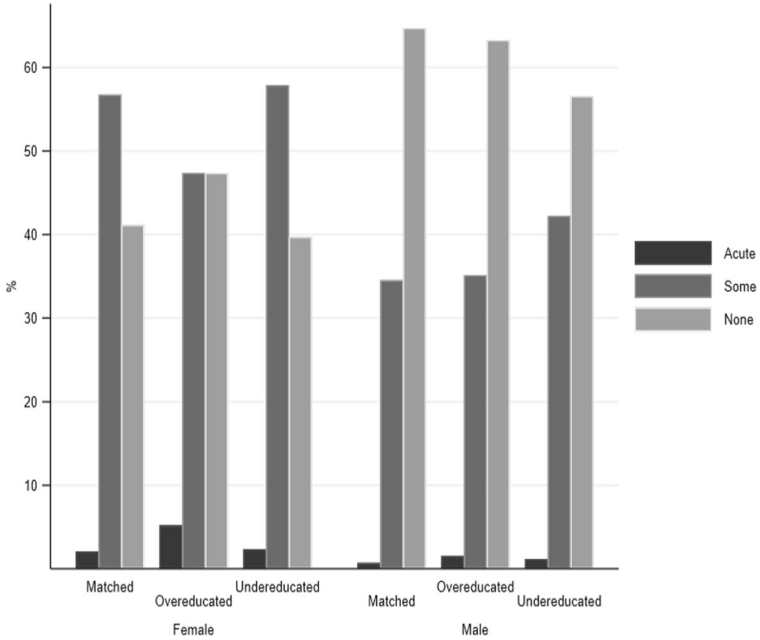

Fig. 2 Distribution of self-reports on depression over categories of educational mismatch and gender

"some" ("none") category of anxiety/depression more (less) frequently than their matched counterparts.[6]

3.2 Estimation Strategy

As the baseline specification, we estimate ordered probit models for the selected EQ-5D metrics of pain and anxiety/depression. Our analysis is stratified by gender as has been done in the earlier literature (Campos-Serna et al., 2013; Cottini, 2012; Robone et al., 2011). Since our health outcomes are self-reported measures, instead of the latent levels of pain and anxiety/depression H^*_{it} we can only observe an indicator of the category H_{it} to which our latent variables belong:

$$H_{\mathrm{it}} = \mathrm{j} \text{ if } \mu_{\mathrm{j-1}} < H^*_{\mathrm{it}} < \mu_{\mathrm{j}},\ \mathrm{j} = 1 \ldots 2, \text{ where } \mu_0 = -\infty, \mu_{\mathrm{j-1}} \le \mu_{\mathrm{j}}, \mu_2 = +\infty \tag{1}$$

Hence, the latent variables H^*_{it} can be represented as a linear function of a vector of the main explanatory variables, socioeconomic and work-related covariates, and a random error term ε_{it} which is assumed to be normally distributed. However, self-reported outcomes measured on a categorical scale (e.g., self-assessed health, EQ-5D metrics, life and job satisfaction, health system responsiveness etc.) tend to suffer from the issue of reporting heterogeneity, also known as a differential item functioning (DIF) (Bzostek et al., 2016; Knott et al., 2017; Rice et al., 2012; Rossouw et al., 2018). This problem arises from the

[6] Along with the graphical representation, conclusions about statistical significance of differences are done on the basis of t-test of equal means (not reported here, but available upon request).

fact that individuals vary in terms of understanding and using ordinal response categories. As a result, they place the cut points between adjacent response categories in different ways (Grol-Prokopczyk et al., 2015; King & Wand, 2007).

The ordered choice models fail to address the issue of reporting heterogeneity and, consequently, may provide biased results. Thus, as the next step, we test the validity of our baseline specifications by estimating an extension of the standard OPROBIT model—the hierarchical ordered probit (HOPIT) model. The HOPIT model allows adjustment for the reporting heterogeneity bias by disentangling the total effect of an explanatory variable on a dependent variable into the direct one and the one related to the reporting styles of individuals (Tandon et al., 2003). Along with the dependent variable (self-reported outcome measured on a categorical scale), the HOPIT model requires the use of *anchoring vignettes*—a set of questions which refer to a *hypothetical* individual and situation and portray the *same* outcome measured on the *same* categorical scale as the original dependent variable they complement (O'Doherty et al., 2017; van Soest et al., 2011).

The HOPIT model "works properly" when vignettes fulfil two assumptions—*response consistency* and *vignette equivalence*. *Response consistency* requires individuals to use reporting scales in the same way while assessing vignettes and self-reported outcomes. *Vignette equivalence* states that "respondents may differ with each other in how they perceive the level of the variable portrayed in each vignette, but any differences must be random and hence independent of the characteristic being measured" (King et al., 2004, p. 194).

The general specification of the HOPIT model can be written as follows:

$$H_{ik}^{v*} = \eta_k + \varepsilon_{ik.}^{v} \varepsilon_{ik}^{v} \sim N\left(0, \sigma_v^2\right) \tag{2}$$

where H_{ik}^{v*} is a health outcome for vignette k, perceived by individual i and unobserved to the researcher*;* η_k is the mean of the underlying scale for the vignette k; ε_{ik}^{v} is an idiosyncratic error term.

$$h_{ik}^{v} = j, if\, \mu_i^{j-1} \preccurlyeq H_{ik}^{v*} \prec \mu_i^j, \tag{3}$$

with $\mu_i^0 = -\infty$ and $\mu_i^j = \infty$; where h_{ik}^{v} is the vignette rating on a j-point categorical scale observed instead of H_{ik}^{v*}.

$$\mu_i^1 = X_i \gamma^1 + u_i,$$

$$\mu_i^j = \mu_i^{j-1} + exp(X_i \gamma^j) \tag{4}$$

where $u_i \sim N\left(0, \sigma_u^2\right)$ is an unobserved individual specific random effect (assumed to be independent of X_i and the other error terms in the model); γ^j and σ_u^2 are other parameters to be estimated. Hence, the OPROBIT model implies the cut points to be represented by fixed constants (which are common to all individuals), while the HOPIT model treats them as functions of covariates X. The HOPIT model arises from the OPROBIT setting and, thus, can also be estimated by the maximum likelihood techniques. However, its likelihood function contains an additional component which is responsible for the vignettes along with the one related to the self-reported outcome (as in the standard OPROBIT model).

Since the RLMS-HSE questionnaires do not contain vignettes, we adopt the procedure of the HOPIT estimation with externally collected vignettes introduced by Harris et al. (2020). As a source of externally collected vignettes, we use the Russian profile of the

Table 4 Examples of vignette questions used in the WHS

EQ-5D metric: pain		
Vignette 1: Mark has pain in his knees, elbows, wrists and fingers, and the pain is present almost all the time. It gets worse during the first half of the day. Although medication helps, he feels uncomfortable when moving around, holding and lifting things. Overall in the last 30 days, how much of bodily aches or pains did Mark have?		
1—acute	2—some	3—none
EQ-5D metric: anxiety/depression		
Vignette 1: Roberta feels depressed most of the time. She weeps frequently and feels hopeless about the future. She feels that she has become a burden on others and that she would be better dead. In the last 30 days, how much of a problem did Roberta have with worry or anxiety/feeling sad, low or depressed?		
1—acute	2—some	3—none

World Health Survey (WHS). The description of some illustrative vignettes for pain and depression used in the analysis is provided in Table 4. We merge our principal RLMS-HSE (2005) sample with vignettes for pain and anxiety/depression from the WHS (2003) on the basis of such sociodemographic characteristics as age, gender, marital status, educational attainment and income.[7] The respondents from the principal and the vignette samples (obtained through merging) reveal statistically significant differences in terms of sociodemographic and work-related characteristics (results available upon requests).[8] In order to adjust for this disparity, we construct post-stratification weights and assign them to the vignette sample, as proposed by Harris et al. (2020). As such correction is implemented, the vignette sample does not differ from the principal in terms of means any longer (results available upon request).[9]

The self-reported measures of pain and anxiety/depression are defined on a 3-point categorical scale and increase from the worst to the best health outcome in our analysis. The original vignettes in the WHS are measured on a 5-point categorical scale and take the values of "none", "mild", "moderate", "severe", and "extreme" outcomes. We provide correspondence between our dependent variables and associated vignettes by rescaling the latter from the worst to the best health outcome. We also reduce the number of categories from 5 to 3 by keeping the "none" category as given and merging "mild" and "moderate" categories under the "some" outcome, and "severe" and "extreme" categories under the "acute" outcome.

4 Results

The estimated coefficients and standard errors from the OPROBIT and HOPIT models for pain and anxiety/depression (stratified by gender) are displayed in Tables 5 and 6. To identify the parameters of an OPROBIT model it is customary to fix the variance to 1 (for example see Jones et al. (2007)), while the variance in the HOPIT model is an estimated parameter. Therefore, to make the estimated coefficients of the two models comparable, in Table 5 and 6 the reported coefficients in the OPROBIT are equal to the coefficients from the "original OPROBIT" multiplied by the variance in the HOPIT model (see Jones et al. (2007), p. 70). Table 5 provides evidence that the results we get through the OPROBIT model do not change much after the adjustment for reporting heterogeneity through the

[7] The following age groups "18–22", "23–27", "28–32", "33–37", "38–42", "43–47", "48–52", "53–57", "58+" are specified for the purpose of merging. Gender and marital status are taken as given. Educational attainment is constructed on the basis of years of education due to the lack of consistency between the categorical measures of education in the RLMS-HSE and the WHS datasets. Respondents are split in the following groups "8–11", "12–15", "16–19", "20+" with respect to the number of years of completed education. Both the RLMS-HSE and the WHS datasets contain permanent income indicators. This allows us to construct income quintiles in both datasets by means of using similar indicators (number of rooms, car, TV, VCR, DVD, washing machine, fridge, fixed-line phone, computer, other apartment) in the principal component analysis.

[8] We follow the parametric method of estimation of the HOPIT model which does not require the number of observations for the vignette responses to be equivalent to the total number of observations in the principal sample. However, the vignette sample should not reveal any statistically significant differences in terms of means with the principal sample. If any occur, they should be corrected by means of applying post-stratification weights.

[9] Once post-stratification weights are defined, they need to be applied to the vignette component of the HOPIT likelihood function (Harris et al., 2020).

Table 5 Coefficients and standard errors of the independent variables in the OPROBIT and HOPIT model: results for *Pain*

	Women (n = 1,732)				Men (n = 1,658)			
	Ordered probit		Weighted HOPIT		Ordered probit		Weighted HOPIT	
	Coef	St. Err	Coef	St. Err	Coef	St. Err	Coef	St. Err
Age	−0.543***	0.033	−0.629***	0.077	−0.544***	0.033	−0.474***	0.098
Age^2	0.090	0.035	0.464***	0.085	−0.025	0.029	−0.118	0.083
Single	−0.003	0.064	0.225	0.144	−0.028	0.085	−0.375	0.251
Education	0.051	0.066	0.324**	0.148	0.225*	0.069	0.322*	0.193
Low income group	0.043	0.092	−0.036	0.204	−0.121	0.097	−0.925***	0.277
Medium low income group	0.077	0.091	0.376*	0.208	−0.138	0.096	−0.806***	0.263
Medium high income group	0.249***	0.091	−0.096	0.202	−0.099	0.096	−0.977***	0.277
High income group	0.166*	0.093	0.319	0.216	0.089	0.100	−0.622**	0.282
Frequent visits to GP (12 m)	−0.743***	0.059	−0.687***	0.131	−1.112***	0.071	−1.220***	0.198
Metropolitan area	0.547***	0.091	0.463*	0.204	0.442**	0.094	0.399	0.270
Large company	−0.123	0.071	−0.256	0.160	0.273*	0.072	0.334	0.206
Public sector	0.040	0.062	0.065	0.137	0.494***	0.062	0.455***	0.174
Occupational change (12 m)	−0.215	0.088	−0.176	0.192	−0.111	0.076	−0.168	0.218
Unemployment (12 m)	0.015	0.146	0.008	0.327	0.368	0.187	0.765	0.573
Overeducation	−0.159	0.089	−0.211	0.198	−0.049	0.090	−0.284	0.258
Undereducation	−0.180*	0.079	−0.318*	0.171	−0.201	0.086	0.170	0.236
Cut point Eq. 1								
Age			0.066	0.042			0.343***	0.043
Age^2			0.071	0.047			−0.137***	0.040
Single			0.203**	0.085			−0.180	0.118
Education			−0.071	0.085			−0.148*	0.083
Low income group			−0.098	0.119			−0.087	0.126
Medium low income group			−0.238**	0.117			−0.381***	0.125
Medium high income group			0.033	0.124			−0.303***	0.124

Table 5 (continued)

	Women (n = 1,732)				Men (n = 1,658)			
	Ordered probit		Weighted HOPIT		Ordered probit		Weighted HOPIT	
	Coef	St. Err	Coef	St. Err	Coef	St. Err	Coef	St. Err
High income group			−0.156	0.127			0.291***	0.122
Frequent visits to GP (12 m)			−0.059	0.074			−0.201***	0.090
Metropolitan area			−0.031	0.114			0.085	0.117
Large company			0.026	0.094			−0.239	0.087
Public sector			−0.038	0.078			0.147	0.077
Occupational change (12 m)			0.158	0.109			−0.086	0.097
Unemployment (12 m)			0.356**	0.175			0.764	0.236
Overeducation			0.398***	0.114			0.109	0.107
Undereducation			0.127	0.103			0.138	0.111
Cut point Eq. 2								
Age			−0.081	0.058			0.042	0.080
Age^2			0.431***	0.071			−0.096	0.068
Single			0.221*	0.114			−0.366	0.206
Education			0.356***	0.117			0.123	0.152
Low income group			−0.073	0.163			−0.907	0.226
Medium low income group			0.433***	0.166			−0.699	0.201
Medium high income group			−0.439***	0.163			−0.950	0.226
High income group			0.231	0.176			−0.840	0.230
Frequent visits to GP (12 m)			0.089	0.102			−0.096	0.157
Metropolitan area			−0.097	0.157			−0.057	0.219
Large company			−0.177	0.129			0.103	0.163
Public sector			0.046	0.107			−0.051	0.140
Occupational change (12 m)			0.021	0.147			−0.029	0.177

Table 5 (continued)

	Women (n = 1,732)				Men (n = 1,658)			
	Ordered probit		Weighted HOPIT		Ordered probit		Weighted HOPIT	
	Coef	St. Err	Coef	St. Err	Coef	St. Err	Coef	St. Err
Unemployment (12 m)			−0.068	0.239			0.381	0.470
Overeducation			−0.154	0.157			−0.294	0.212
Undereducation			−0.226*	0.132			0.406**	0.188
Log-likelihood	−1339.692	−2286.035	−1204.769	−2027.199				

The coefficients of the ordered probit models are adjusted for the variance of the HOPIT model. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 6 Coefficients and standard errors of the independent variables in the OPROBIT and HOPIT model: results for *Anxiety/Depression*

	Women (n = 1,732)				Men (n = 1,658)			
	Ordered probit		Weighted HOPIT		Ordered probit		Weighted HOPIT	
	Coef	St. Err	Coef	St. Err	Coef	St. Err	Coef	St. Err
Age	−0.221***	0.032	−0.230***	0.063	−0.356***	0.034	−0.686***	0.077
Age^2	−0.003	0.035	0.013	0.070	0.054	0.030	0.012	0.065
Single	−0.109*	0.064	−0.489***	0.129	−0.325**	0.087	−1.204***	0.187
Education	0.025	0.066	0.274**	0.131	0.192*	0.071	−0.047	0.149
Low income group	0.018	0.092	0.634***	0.179	−0.128	0.101	0.961***	0.216
Medium low income group	0.040	0.091	0.649***	0.181	−0.209	0.100	0.059	0.213
Medium high income group	0.162	0.091	0.437**	0.175	−0.030	0.100	−0.311	0.202
High income group	0.365*	0.094	0.867***	0.189	−0.247	0.102	0.457**	0.221
Frequent visits to GP (12 m)	−0.334***	0.059	−0.234**	0.115	−0.404**	0.073	−0.196	0.153
Metropolitan area	0.399***	0.091	0.108	0.176	0.307	0.098	0.033	0.199
Large company	0.004	0.072	0.055	0.141	0.123	0.075	0.111	0.155
Public sector	−0.046	0.062	−0.092	0.121	0.202**	0.064	0.253**	0.133
Occupational change (12 m)	−0.063	0.088	−0.113	0.173	−0.286**	0.078	−0.246	0.161
Unemployment (12 m)	0.158	0.144	−0.060	0.295	0.226	0.194	−0.300	0.379
Overeducation	0.001	0.089	0.416**	0.180	−0.120	0.093	−0.353*	0.190
Undereducation	−0.017	0.080	−0.145	0.155	−0.260**	0.088	−0.099	0.187
Cut point Eq. 1								
Age			0.066	0.044			0.116**	0.048
Age^2			−0.018	0.049			−0.235***	0.045
Single			0.306***	0.093			−0.112	0.126
Education			−0.062	0.093			−0.017	0.093
Low income group			0.151	0.123			−0.091	0.136
Medium low income group			−0.184	0.126			−0.468***	0.131
Medium high income group			0.385***	0.130			0.145	0.123

Table 6 (continued)

	Women (n = 1,732)				Men (n = 1,658)			
	Ordered probit		Weighted HOPIT		Ordered probit		Weighted HOPIT	
	Coef	St. Err	Coef	St. Err	Coef	St. Err	Coef	St. Err
High income group			0.157	0.132			−0.303**	0.135
Frequent visits to GP (12 m)			0.090	0.080			0.092	0.096
Metropolitan area			−0.029	0.124			−0.010	0.125
Large company			−0.013	0.098			0.027	0.096
Public sector			−0.008	0.085			−0.112	0.083
Occupational change (12 m)			0.026	0.122			0.115	0.103
Unemployment (12 m)			0.313	0.215			0.105	0.232
Overeducation			−0.057	0.115			0.162	0.122
Undereducation			−0.132	0.112			0.200*	0.123
Cut point Eq. 2								
Age			−0.019	0.046			−0.357***	0.058
Age^2			0.020	0.051			−0.034	0.052
Single			−0.510***	0.096			−0.923***	0.145
Education			0.291***	0.094			−0.263*	0.113
Low income group			0.717***	0.129			1.164***	0.166
Medium low income group			0.763***	0.132			0.330**	0.165
Medium high income group			0.276**	0.124			−0.307**	0.151
High income group			0.576***	0.136			0.784***	0.170
Frequent visits to GP (12 m)			0.108	0.082			0.208*	0.114
Metropolitan area			−0.313**	0.123			−0.284*	0.147
Large company			0.060	0.100			−0.011	0.116
Public sector			−0.052	0.086			0.061	0.099
Occupational change (12 m)			−0.063	0.123			0.039	0.122

Table 6 (continued)

	Women (n = 1,732)				Men (n = 1,658)			
	Ordered probit		Weighted HOPIT		Ordered probit		Weighted HOPIT	
	Coef	St. Err	Coef	St. Err	Coef	St. Err	Coef	St. Err
Unemployment (12 m)			− 0.269	0.219			− 0.563**	0.264
Overeducation			0.471***	0.131			− 0.268*	0.143
Undereducation			− 0.147	0.109			0.160	0.143
Log-likelihood	− 1326.129	− 2331.537	− 1113.257	− 1994.719				

The coefficients of the ordered probit models are adjusted for the variance of the HOPIT model. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

HOPIT model, since the magnitude and level of statistical significance of the coefficients for overeducation and undereducation in the main equation of the two models are similar. Such coefficients are negative and statistically significant only for undereducation of women, suggesting that undereducated women experience more pain than their matched counterparts. When considering the influence of undereducation on pain for women the magnitude (in absolute terms) of the coefficient is smaller (in absolute terms) in the OPROBIT than in the HOPIT model (0.180 and 0.318, respectively). This disparity between the models is explained by the presence of a “more optimistic” reporting style of undereducated Russian women. Indeed, a statistically significant coefficient of −0.226 in the second cut point equation of the HOPIT model indicates that when choosing between “some” and “none” categories of pain, undereducated women are more likely to report the absence of pain than their matched counterparts. Differently, undereducated men tend to adopt a more “pessimistic” reporting style (indicated by a positive and statistically significant coefficient of 0.406 in the second cut point equation of the HOPIT model); this result seems to be consistent with the fact that the coefficient for this variable in the main equation changes from being negative in the OPROBIT to be positive in the HOPIT model, although such coefficient is not statistically significant.

Table 6 shows a positive influence of overeducation on anxiety/depression for women (indicated by a statistically significant coefficient of 0.416 in the main equation of the HOPIT model). However, this effect is disguised in the OPROBIT model by the “pessimistic” reporting style of Russian women (indicated by a statistically significant coefficient of 0.471 in the second cut point equation of the HOPIT model). Undereducation does not seem to influence depression in the female sub-sample (either directly or through reporting styles). Overeducated men in Russia tend to be worse off in terms of depression than their matched counterparts (indicated by a negative statistically significant coefficient of −0.353 in the main equation of the HOPIT model). However, this effect “is hidden” in the OPROBIT model (where the coefficient is not statistically significant) due to the optimistic reporting style of men (shown by a statistically significant coefficient of −0.268 in the second cut point equation). The negative influence of undereducation on depression among Russian men observed in the OPROBIT model seems to disappear after adjusting for reporting heterogeneity (indeed the coefficients for undereducation in the cut point equations are positive and one of them is statistically significant).

Overall, our results provide evidence that educational mismatch is related to the health status of Russian employees (either directly or through reporting styles), but these associations appear more marked for depression than pain. Moreover, the direction and the magnitude of these effects are gender specific. For instance, when computing marginal effects (results are available upon request) overeducation appears to increase the probability of not being depressed by 15.4 percentage points among Russian women, and to increase the probability of being depressed by 11.6 percentage points among Russian men. Thus, our finding for the male employees are in line with the earlier studies indicating a negative impact of overeducation on mental health and psychological well-being of employees in the European countries (Atres et al., 2013; Bracke, 2014; Garcy, 2015; Salinas-Jimenez et al., 2016; Ueno & Krause, 2018).

5 Discussion

Our study contributes to the literature on socioeconomic determinants of health and investigates the influence of vertical educational mismatch, that is the divergence of the level of education an individual has from the one required for the position he/she holds in the labour market, using selected EQ-5D metrics of the Russian working population—pain and anxiety/depression. Since our dependent variables are self-reported and measured on a categorical scale, along with obtaining the overall effect from the ordered probit models, we also test for the presence of reporting heterogeneity bias in our estimates. To address this issue, we adopt the procedure of estimating the HOPIT model with externally collected vignettes, as proposed by Harris et al. (2020).

We believe that our analysis sheds some light on the mechanisms which stand behind the influence of educational mismatch on self-reported health outcomes of employees, for both the physical and psychological components. Our findings provide evidence that, after adjusting for reporting heterogeneity, undereducation has a negative influence on the physical component of health (proxied by pain) for women, while overeducation affects the psychological one (proxied by anxiety/depression) in both gender groups. Overeducated women appear to have better psychological health, while overeducated men worst psychological health, than their matched counterparts. More specifically, overeducation appears to increase the probability of not being depressed by about 15 percentage points among women, while it increases the probability of being depressed by about 11 percentage points among men. In addition, both types of vertical educational mismatch affect the perception of health the individuals express through their reporting styles. When considering depression, the negative influence of undereducation among men observed in the ordered probit model seems to disappear after adjusting for reporting heterogeneity. Differently, the positive (negative) influence of overeducation on anxiety/depression for women (men) showed by the results of the HOPIT model is disguised in the ordered probit model by the “pessimistic” (“optimistic”) reporting style of women (men).

Our analysis has some limitations and possible ways of extension. Given the cross-sectional nature of the dataset we use, in our analysis we are not able to claim that the relationship between educational mismatch and health is truly causal. One way to address the issue of endogeneity would be by referral to a panel dataset; however, unfortunately data about vignettes which are present in the WHS and can be matched with RLMS-HSE are only available as a cross-section. Another possibility for addressing the issue of endogeneity would be to adopt an instrumental variable approach, referring to some variables to instrument the educational mismatch dummies (e.g. relevant macro indicators which describe the current state of the labour market and educational system in Russia). However, to the best of our knowledge, the HOPIT model, on which we heavily rely in our study, has not been extended yet to allow for the inclusion of instrumental variables in the model. Therefore, even the possibility of using instrumental variables does not appear doable in our case.

Although our HOPIT and ordered probit models seem to be consistent, the use of distinct time periods for selected EQ-5D metrics and associated vignettes—RLMS-HSE (2005) and World Health Survey (2003)—tends to challenge the response consistency assumption. However, we assume that no unexpected socioeconomic and political shocks took place in Russia between 2003 and 2005 which would drastically change the reporting styles of the population. As a solution, we could replicate our analysis in the future by using internal country-specific vignettes for pain and anxiety if they are made available in the RLMS-HSE. In addition, we could conduct a comparative cross-country analysis by taking into consideration other European (both Western and Eastern) and Asian countries. Finally, we could consider in more detail the psychological mechanisms which stand behind the gender differences in employees' self-perception of health (proxied by their reporting styles) in relation to educational mismatch.

From the policy point of view, we would suggest focusing on reducing the number of overeducated male employees in the labor market. This goal can be achieved by making blue collar occupations more attractive for employment by means of improving working conditions and providing competitive rewards and fringe benefits. In the long run, this may nudge male high school graduates to choose employment in these occupations instead of applying for higher professional education and ending up in inefficient utilization of human capital: overeducation. In order to reduce the number of undereducated female employees in the economy, we would recommend policy makers put an emphasis on promoting and financing adult vocational and professional training aimed at closing the knowledge gap in the female segment of the Russian labour market. In addition, undereducated women need to be included in mentoring programs in the workplace. This might help to reduce stress caused by assigning challenging tasks to undereducated women which they are not properly qualified to do, thus avoiding the risk of a heavy workload and professional burnout, and to keep the work-life balance which is particularly important for women in Russia. Overall, the suggested policy implications might help to overcome both the negative impact of inefficient utilization of human capital in the Russian economy and the negative spillovers of educational mismatch on health and well-being of Russian employees.

Acknowledgements We are grateful for their suggestions and comments to Olena Nizalova, Luca Corazzini, Giuseppe Porro, Charles Becker, Bruce Hollingsworth and the participants to XXXIV Conference of the Italian Society of Public Economics, LV SIEDS Conference and the Economics Department Seminars of the University of Navarra, Pamplona. We are very grateful to Maria Vasiakina for the research work she has conducted on an early version of the paper, that was used for her PhD Dissertation. The usual disclaimers apply.

Funding Open access funding provided by Università degli Studi del Piemonte Orientale Amedeo Avogrado within the CRUI-CARE Agreement. No funding was received to assist with the preparation of this manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Acosta-Ballesteros, J., Osorno-del Rosal, M. P., & Rodriguez-Rodriguez, O. M. (2017). Overeducation of young workers in Spain: How much does the first job matter? *Social Indicators Research*. https://doi.org/10.1007/s11205-017-1643-z

Artes, J., Salinas-Jimenez, M. M., & Salinas-Jimenez, J. (2013). Small fish in a big pond or big fish in a small pond? The effects of educational mismatch on subjective wellbeing. *Social Indicators Research*. https://doi.org/10.1007/s11205-013-0526-1

Badillo-Amador, L., & Vila, L. E. (2013). Education and skill mismatches: Wage and job satisfaction consequences. *International Journal of Manpower, 34*(5), 416–428.

Baert, S., Cockx, B., & Verhaest, D. (2013). Overeducation at the start of the career: stepping stone or trap? *Labour Economics, 25*, 123–140.

Becker, G. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). University of Chicago Press.

Bracke, P., van de Straat, V., & Missinne, S. (2014). Education, mental health, and education-labor market misfit. *Journal of Health and Social Behavior, 55*(4), 442–459.

Bzostek, S., Sastry, N., Goldman, N., Pebley, A., & Duffy, D. (2016). Using vignettes to rethink Latino-white disparities in self-rated health. *Social Science and Medicine, 149*, 46–65.

Campos-Serna, J., Ronda-Perez, E., Artazcoz, L., Moen, B. E., & Benavides, F. G. (2013). Gender inequalities in occupational health related to the unequal distribution of working and employment conditions: A systematic review. *International Journal for Equity in Health, 2013*(12), 1–18.

Cottini, E. (2012). Is your job bad for your health? Explaining differences in health at work across gender. *International Journal of Manpower, 33*(3), 301–321.

Davia, M. A., McGuinness, S., & O'Connell, P. J. (2017). Determinants of regional differences in rates of overeducation in Europe. *Social Science Research, 63*, 67–80.

Diem, A., & Wolter, S. C. (2014). Overeducation among Swiss university graduates: Determinants and consequences. *Journal of Labour Market Research, 47*, 313–328.

Ferguson, B. D., Tandon, A., Gakidou, E., & Murray, C. J. L. (2003). Estimating permanent income using indicator variables. In C. J. L. Murray & D. B. Evans (Eds.), *Health systems performance assessment: Debates, methods and empiricism* (pp. 748–760). World Health Organization.

Garcy, A. M. (2015). Educational mismatch and mortality among native-born workers in Sweden. A 19-year longitudinal study of 2.5 million over-educated, matched and under-educated individuals, 1990–2008. *Sociology of Health & Illness, 37*(8), 1314–1336.

Ghaffarzadegan, N., Xue, Y., & Larson, R. C. (2017). Work-education mismatch: An endogeneity theory of professionalization. *European Journal of Operational Research, 261*, 1085–1097.

Grol-Prokopczyk, H., Verdes-Tennant, E., McEniry, M., & Ispany, M. (2015). Promises and pitfalls of anchoring vignettes in health survey research. *Demography, 52*(5), 1703–1728.

Handel, M.J., Valerio, A., & Sanchez Puerta, M.L. (2016). *Accounting for Mismatch in Low- and Middle-Income Countries: Measurement, Magnitudes, and Explanations. Directions in Development*. Washington, DC: World Bank.

Hardie, J. H. (2014). The consequences of unrealized occupational goals in the transition to adulthood. *Social Science and Medicine, 48*, 196–211.

Harris, M. N., Knott, R. J., Lorgelly, P. K., & Rice, N. (2020). Using externally collected vignettes to account for reporting heterogeneity in survey self-assessment. *Economics Letters*. https://doi.org/10.1016/j.econlet.2020.109325

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of the EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*, 1727–1736.

Hultin, H., Lundberg, M., Lundin, A., & Magnusson, C. (2016). Do overeducated individuals have increased risks of ill health?: A Swedish population-based cohort study. *Sociology of Health & Illness, 38*(6), 980–995.

Iriondo, I., & Perez-Amaral, T. (2016). The effect of educational mismatch on wages in Europe. *Journal of Policy Modeling, 38*, 304–323.

Jones, A.M., Rice, N., Bago d'Uva, T., & Balia, S. (2007). *Applied Health Economics*. Routledge advanced texts in economics and finance.

Kampelmann, S., & Rycx, F. (2012). The impact of educational mismatch on firm productivity: Evidence from linked panel data. *Economics of Education Review, 31*, 918–931.

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.

King, G., Murray, C. J. L., Salomon, S. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*(1), 191–207.

King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis, 15*, 46–66.

Knott, R. J., Longelly, P. K., Black, N., & Hollingsworth, B. (2017). Differential item functioning in quality of life measurement: An analysis using anchoring vignettes. *Social Science and Medicine, 190*, 247–255.

Kupets, O. (2016a). Education-job mismatch in Ukraine: Too many people with tertiary education or too many jobs for low-skilled? *Journal of Comparative Economics, 44*, 125–147.
Kupets, O. (2016b). Skill mismatch and overeducation in transition economies. *IZA World of Labour, 224*, 1–10.
Kyui, N. (2010). Returns to education and education-occupation mismatch within a transition economy. Empirical analysis for the Russian Federation. *Documents de travail du Centre d'Economie de la Sorbonne*, 2010.
Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics, 107*(2), 407–437.
Marmot, M., Allen, J., Bell, R., Bloomer, E., & Goldblatt, P. (2012). WHO European review of social determinants of health and the health divide. *The Lancet, 380*(9846), 1011–1029. https://doi.org/10.1016/s0140-6736(12)61228-8
Mateos-Romero, L., & Salinas-Jimenez, M. M. (2018). Labor mismatches: Effects on wages and job satisfaction in 17 OECD countries. *Social Indicators Research*. https://doi.org/10.1007/s11205-017-1830-y
Mavromaras, K., & McGuinness, S. (2012). Overskilling dynamics and educational pathways. *Economics of Education Review, 31*, 619–628.
McGuinness, S., Pouliakas, K., & Redmond, P. (2017). Skills mismatch: concepts, measurement and policy approaches. *Journal of Economic Surveys*, 1–31.
Meroni, E. C., & Vera-Toscano, E. (2017). The persistence of overeducation among recent graduates. *Labour Economics, 48*, 120–143.
Milner, A., Aitken, Z., Kavanagh, A., LaMontagne, A. D., & Petri, D. (2017). Status inconsistency and mental health: A random effects and instrumental variables analysis using 14 annual waves of cohort data. *Social Science and Medicine, 189*, 129–137.
Morgado, A., Sequeira, T. N., Santos, M., Ferreira-Lopes, A., & Reis, A. B. (2016). Measuring labour mismatch in Europe. *Social Indicators Research, 129*, 161–179.
Mossakowski, K. N. (2011). Unfulfilled expectations and symptoms of depression among young adults. *Social Science and Medicine, 73*, 729–736.
O'Doherty, M. G., French, D., Steptoe, A., & Kee, F. (2017). Social capital, deprivation and self-rated health: Does reporting heterogeneity play a role? Results from the English Longitudinal Study of Ageing. *Social Science and Medicine, 179*, 191–200.
Pecoraro, M. (2016). The incidence and wage effects of overeducation using the vertical and horizontal mismatch in skills. *International Journal of Manpower, 37*(3), 536–555.
Piper, A. (2015). Heaven knows I'm miserable now: Overeducation and reduced life satisfaction. *Education Economics, 23*, 677–692.
Rabin, R., & de Charro, F. (2001). EQ-5D: A measure of health status from the EuroQol Group. *Annals of Medicine, 335*, 337–343.
Cedefop Report (2018). Insights into skill shortages and skill mismatch: Learning from Cedefop's European Skills and Jobs Survey. Luxembourg: Publications Office of the European Union
Rice, N., Robone, S., & Smith, P. C. (2012). Vignettes and health systems responsiveness in cross-country comparative analyses. *Journal of the Royal Statistical Society, 175*(2), 337–369.
Robone, S. (2023). Insights into Alcohol Consumption: does Educational Mismatch affect Drinking Behaviors in Russia?. *Italian Review of Economics, Demography and Statistics*.
Robone, S., Jones, A. M., & Rice, N. (2011). Contractual conditions, working conditions and their impact on health and well-being. *European Journal of Health Economics, 12*(5), 429–444.
Romer, P. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5), S71–S102.
Romero, L. M., Huertas, I. P. M., & Jimenez, M. M. S. (2017). Wage effects of cognitive skills and educational mismatch in Europe. *Journal of Policy Modeling, 39*, 909–927.
Rossouw, L., Bago d'Uva, T., & van Dooerslaer, E. (2018). Poor health reporting? Using anchoring vignettes to uncover health disparities by wealth and race. *Demography, 55*, 1935–1956.
Salinas-Jimenez, M. M., Artes, J., & Salinas-Jimenez, J. (2016). Educational mismatch and job aspiration. A subjective wellbeing analysis using quantile regression. *International Journal of Manpower, 37*(1), 115–134.
Schultz, T. (1963). *The economic value of education*. Columbia University.
Sellami, S., Verhaest, D., Nonneman, W., & Van Trier, W. (2017). The impact of educational mismatches on wages: the influence of measurement error and unobserved heterogeneity. *The B.E. Journal of Economic Analysis and Policy*. https://doi.org/10.1515/bejeap-2016-0055
Tandon, A., Murray, C. J. L., Salomon, J. A., & King, G. (2003). Statistical models for enhancing cross-population comparability. In C. J. L. Murray & D. B. Evans (Eds.), *Health systems performance assessment: Debates, methods and empiricism* (pp. 727–746). World Health Organization.
Ueno, K., & Krause, A. (2018). Overeducation, perceived career progress, and work satisfaction in young adulthood. *Research in Social Stratification and Mobility, 55*, 51–62.

Van Soest, A., Delaney, L., Harmon, C., Kapteyn, A., & Smith, J. P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society, 174*(3), 575–595.

Verdugo, R., & Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources, 24*(4), 629–643.

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.

Vossemer, J., & Schuck, B. (2015). Better overeducated than unemployed? The Short- and long-term effects of an overeducated labour market re-entry. *European Sociological Review*, 1–15.

Zhu, R., & Chen, L. (2016). Overeducation, overskilling and mental well-being. *BE Journal of Economic Analysis and Policy*. https://doi.org/10.1515/bejeap-2015-0187