Education Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cede20

Educational and skills mismatches: unravelling their effects on wages across Europe

L. Cultrera, B. Mahy, F. Rycx & G. Vermeylen

To cite this article: L. Cultrera, B. Mahy, F. Rycx & G. Vermeylen (2022): Educational and skills mismatches: unravelling their effects on wages across Europe, Education Economics, DOI: 10.1080/09645292.2022.2050995

To link to this article: https://doi.org/10.1080/09645292.2022.2050995

Published online: 09 Mar 2022.

EDUCATION ECONOMICS
https://doi.org/10.1080/09645292.2022.2050995

Educational and skills mismatches: unravelling their effects on wages across Europe

L. Cultrera [a], B. Mahy[b], F. Rycx [c] and G. Vermeylen [d]

[a]Loredana Cultrera - University of Mons (UMONS - Belgium), Soci&ter and Risk, Mons, Belgium; [b]Benoît Mahy – University of Mons (UMONS - Belgium), Soci&ter, and DULBEA, Mons, Belgium; [c]François Rycx - Université Libre de Bruxelles (ULB - Belgium), CEB, DULBEA, IRES, GLO, Soci&ter and IZ, Brussels, Belgium; [d]Guillaume Vermeylen – University of Mons (UMONS - Belgium), Soci&ter, DULBEA and CEB, Mons, Belgium

ABSTRACT
This paper is among the first to investigate the impact of over-education and over-skilling on workers' wages using a unique pan-European database covering twenty-eight countries for the year 2014, namely the CEDEFOP's European Skills and Jobs (ESJ) survey. Overall, the results suggest the existence of a wage penalty associated with over-education. When the educational and the skills mismatches are interacted with each other in order to distinguish apparent over-education from genuine over-education, the results highlight that the workers with the highest wage penalty are those who are both over-educated and over-skilled.

ARTICLE HISTORY
Received 9 February 2021
Accepted 22 February 2022

KEYWORDS
Educational mismatch; skills mismatch; wages; European survey

JEL CODES
J21; J24

1. Introduction

It is undeniable that the level of education in advanced industrial countries has increased significantly in recent years (Barro and Lee 2013). In the last decade, the European Union faced a massive increase in the number of tertiary graduates and has met its target of having at least 40% of the population aged between 30 and 34 attaining tertiary education (European Commission 2019). Indeed, in 2019, the share of 30–34 year-olds having successfully completed tertiary education was evaluated at 40.7%, compared to 32% in 2009. At the same time, the level of requirements for jobs has also risen (Green 2007), leading researchers to investigate how these two evolutions are related to each other (Korpi and Tahlin 2009). If the level of education required by jobs does not rise at the same rate as the level of education attained by workers, educational mismatch may arise (Freeman 1976). This mismatch represents the difference between the level of education of a worker and the level of education required for her/his job. A mismatched worker is considered as over-educated if her/his level of education is higher than that required to perform her/his job, or under-educated if her/his level of education is lower than that required for her/his job. Focussing on the first situation, the European Commission (2017) shows that over-education is an important and growing phenomenon that concerned, on average, 40.2% of workers in the EU28 countries over the years 2002–2016.

Most empirical studies on the impact of over-education on wages consistently find the same two results: (i) over-educated workers suffer a wage penalty compared to equally educated classmates employed in jobs matching their level of education; and (ii) those workers get a wage premium in comparison with their adequately educated colleagues doing the same jobs (Duncan and Hoffman 1981; Cohn and Khan 1995; Sloane, Battu, and Belfield 1999; Allen and van der Velden

CONTACT G. Vermeylen Guillaume.Vermeylen@umons.ac.be

2001; Bauer 2002; McGuinness 2006; Lindley and McIntosh 2010; Leuven and Oosterbeek 2011; Bender and Heywood 2011).

Nevertheless, over-education may not reflect workers' real job competencies, and any current evidence that addresses the effect of over-education on wages as such may therefore be biased. Indeed, over-educated workers may possess the required level of skills needed to perform their jobs, and over-education does not necessarily imply over-skilling (Pellizzari and Fichen 2013; Pecoraro 2014; McGuinness, Pouliakas, and Redmond 2018). Given that workers' skills may also differ when considering informal skills acquired through experience, on-the-job training, or even innate abilities (Chevalier 2003; Verhaest and Omey 2009; Chevalier and Lindley 2009), one solution to mitigate such bias consists in refining the definition of educational mismatch by making the educational and skills mismatches interact with each other. Some empirical studies have proposed alternative measurements in order to better account for the heterogeneity among workers and jobs[1], with a common main result that the wage impact of over-education is overestimated in studies that assume equally educated workers to be homogenous in their human capital endowment (Chevalier 2003; Chevalier and Lindley 2009; Green and Zhu 2010; Mavromaras et al. 2013; Pecoraro 2014; Pecoraro 2016; Caroleo and Pastore 2018). Put differently, over-education might be associated with other forms of (in)formal over-skilling.

Theoretical studies therefore address the interaction between the mismatch in education and the mismatch in skills (Groot and Maassen van den Brink 2000). For instance, workers may decide to improve their level of schooling in order to compensate for a lack in one of the (other) human capital characteristics, such as ability or experience (Sicherman 1991; Groot 1993). Human capital theory suggests that over-education or over-skilling should arise in the short run only, being the result of imperfect information. In the longer run, adjustments should be made by the employer in order to use available skills fully, or by the employee seeking a better match in order to fulfil her/his productivity potential and to maximize her/his earnings (OECD 2011).

However, as Quintini (2011) shows, other theories based on human capital suggest that both educational and skills mismatches may remain in the longer run. For example, firms that are constantly facing technological changes may find it appropriate to hire workers with a higher level of education in order to accompany these changes while preventing additional costs, which results in over-education in these firms. More globally, they may also have hiring practices that favour over-skilling in order to ensure quick reactions in changing and uncertain environments. Over-education and over-skilling might then serve as complements in the changing production process.

In this technological change framework, an accurate distinction between educational mismatch and skills mismatch may also be necessary. It may well be the case that technological advances actually cause some workers to 'become' under-educated, but those workers may not necessarily be under-skilled if on-the-job training upgrades their skills while their level of education remains unchanged.

Another explanation comes from inefficient job search, whereby workers accept jobs for which they are over- (under-) educated and/or over- (under-) skilled. In that case, candidates do not have all the information about the job for which they are applying and may thus accept jobs with tasks that do not match their level of education or skills.

All in all, these insights show that the impact of over-education has to be considered jointly and in interaction with (other forms of) over-skilling.

[1]Note that some authors decide to investigate over-skilling *besides* over-education, separately. These studies globally show that both over-education and over-skilling are associated with wage penalties, but also that the wage penalty associated with over-education is higher than that associated with over-skilling (Allen and van der Velden 2001; Mavromaras, McGuinness, and Fok 2009; McGuinness and Byrne 2014). They also find that over-skilling matters the most in the workers' satisfaction determination or in their propensity to implement on-the-job search (Allen and van der Velden 2001; Mateos-Romero and Salinas-Joménez 2018), and that severely over-skilled migrant workers (whose language differs from that/those of the host country) suffer a higher pay penalty (Mavromaras, McGuinness, and Fok 2009). McGuinness and Byrne (2014) focus on immigration and, using data on fifteen European countries, show that higher rates of wage penalties are found for over-skilled male migrants, whereas over-education matters the most for female migrants.

To this end, some more recent works refine the definition of educational mismatch through the implementation of alternative measurements of mismatches in terms of education and skills. Some authors consider these variables separately,whereas other studies make the educational and skills mismatches interact with each other, resulting in three specific situations: (i) apparent matching, *i.e.* when a worker is found to be properly educated but over-skilled; (ii) apparent over-education, *i.e.* when a worker is found to be properly skilled but over-educated; and (iii) genuine over-education, *i.e.* when a worker is both over-educated and over-skilled. These studies consistently find that being apparently over-educated leads to a lower wage penalty than being genuinely over-educated, but also that being apparently matched leads to a lower wage penalty than being genuinely over-educated (Green and Zhu 2010; Mavromaras et al. 2013; Pecoraro 2014; Caroleo and Pastore 2018). As such, these recent developments investigating the interacting impact of the educational and skills mismatches consider binary measurements of both mismatches, in other words whether each mismatch exists or not.

In order to add to the existing literature, this paper focuses on the impact of over-education on wages and investigates whether *interactions* between over-education and (potential) over-skilling mitigate the historically known wage penalties at a European level. We therefore postulate, following the original developments of Green, McIntosh, and Vignoles (1999) and Chevalier (2003)[2], that over-education is then either apparent or genuine depending on whether over-educated workers are properly skilled or over-skilled.

Our main contributions are fourfold. First, by statistically computing the number of years of over-education, apparent over-education, and genuine over-education, we go one step further in measuring the impact of the magnitude, not just the existence, of over-education (in interaction with over-skilling) on wages. Second, we take advantage of a unique pan-European dataset covering all EU28 member states, i.e. the European Skills and Jobs (ESJ) survey, and estimate the aforementioned relationships at the European level, whereas other studies consider the relationships for specific countries or subpopulations of workers, such as university graduates. Third, by relying on an integrated dataset, we use information on over-education and over-skilling coming from the same source, i.e. the worker herself/himself, whereas other studies sometimes merge information regarding these variables under interest from different sources, which confers less reliability as such. Fourth, we take into account the potential endogeneity in the relationships and biased estimates by implementing, besides the more classical ordinary least squares (OLS), a two-stage least squares (2SLS) estimation strategy, which has, to our knowledge, never been used before in this context, when estimating the wage impact of over-education in interaction with potential over-skilling.

The rest of this paper is structured as follows. Section 2 lays out the measurement methods, with their advantages and shortcomings, as well as our wage equation model, and Section 3 describes the dataset. The results are presented in Section 4, and Section 5 finally concludes.

2. Methods

Concerning educational mismatch, three main measurement methods can be used to evaluate the matching between the level of education of a worker and that required for her/his job: the job analysis method, the realized matches method, and the worker self-assessment method. These methods allow, in turn, to evaluate the incidence of over- and under-education (Mahy, Rycx, and Vermeylen 2015). First, the job analysis method can be used to evaluate the required level of education for a given job by using, for example, the U.S. Dictionary of Occupational Titles (DOT). This reference

[2]The first paper dealing with the term 'genuine over-education' is that of Green, McIntosh, and Vignoles (1999), who state that the impact of over-education should be evaluated in a more global perspective including the workers' skills endowment. However, the first paper dealing with both apparent and genuine over-education in a wage penalty perspective is that of Chevalier (2003), who finds that apparent (genuine) over-education is associated with a wage penalty ranging between 5% and 11% (22% and 26%).

provides the required level of education by occupation. This objective evaluation method is used, for example, by Rumberger (1987). It may seem very attractive because it relies on explicit and objective definitions and measurements. However, carrying it out on a large scale requires careful and time-consuming work (Hartog 2000). Moreover, it is criticized by Verdugo and Verdugo (1992), who state that the DOT is sometimes based on a single job analyst discussing requirements with the employer, leading to some doubts over the reliability and the validity of this measurement. Second, the realized matches method is based on the educational attainment of workers in each range of occupation. Two alternative measurements can be used. On the one hand, the mean level of education across a range of occupations is calculated, and workers whose educational attainment is greater (lower) than one standard deviation above the mean value for their occupation are considered as over- (under-) educated (Verdugo and Verdugo 1989). On the other hand, the mode of years of education in each worker's occupation is calculated, and workers whose educational attainment is greater (lower) than the mode are considered as over- (under-) educated (Cohn and Khan 1995; Kiker, Santos, and Mendes de Oliveira 1997). However, the main shortcoming of this realized matches method is that it does not measure the real requirements for a job, but rather the actual assignment practice as determined by hiring standards and labour market conditions (Hartog 2000). Finally, survey techniques can be used whereby respondents are directly asked to specify the minimum level of education they consider to be needed to get their job. This worker self-assessment method is used, for example, by Duncan and Hoffman (1981), Sicherman (1991), and, more recently, Allen and van der Velden (2001). It is of particular interest because it allows to gather up-to-date information, and the obtained required level of education corresponds precisely to the respondent's job rather than to an aggregate measurement. However, this method does not rely on rigorous measurement as respondents can overstate the requirements of their own jobs. They can also reproduce hiring standards over the years, leading to major issues in case of a constant increase in the effective workers' level of education over time (Hartog 2000). So far, there has not been any single perfect indicator, as each measurement method has its advantages and shortcomings. The choice of one method over another is therefore mainly driven by data availability.

The workers in our dataset were asked to self-assess the level of education needed to get their jobs. Comparing this level of education with the highest level of education attained by each worker, we can then determine whether respondents are working above or below their own level of education. Two specific questions asked to respondents during the interviews allow us to define a worker as over-educated:

- What is the highest level of education or training (ISCED_Qualification) that you have completed?
- What is the level of qualification (ISCED_Qualification) needed to *get* your job?

A worker is then defined as over-educated if her/his level of attained education is above the level of education this worker considers to be required *to get* her/his job.

Next, the measurement methods for skills mismatch are most often based on workers' self-assessment, whereby respondents are directly asked how they use their skills in their job (Mavromaras, McGuinness, and Wooden 2007). This self-assessment may concern the overall skills required for the job (*'all my skills match with my job'*) or more specific skills (*'I have the numeracy skills needed to perform my job'*). In our dataset, two questions allow us to define a worker as over-skilled:

- What is the highest level of education or training (ISCED_Qualification) that you have completed?
- What is the level of qualification (ISCED_Qualification) needed to *do* your job?

A worker is then defined as over-skilled if her/his level of qualification is above the one required *to do* (*i.e.* to perform) her/his job.

This *modus operandi* has also been followed by other researchers, such as Dolton and Silles (2008) for UK graduates, or more recently by Caroleo and Pastore (2018) for Italian graduates. Other measurements of skills mismatch are presented in the ESJ survey and may be seen, as such, as more accurate than over-education *to do*. However, our aim was to deepen the previous analysis of the impact of skills mismatch on wages by considering the magnitude of the mismatch phenomenon. For this purpose, we proxy years of over-skilling by years of over-education *to do the worker's job*. We were not able to proxy years of skilling through other variables available in the ESJ dataset. We should highlight, however, that individuals considered as over-educated *to do* also appear to be considered as over-skilled when relying on other ESJ survey questions. In other words, it seems that over-education *to do* is an appropriate proxy for over-skilling.

One originality of this paper lies in our ability to compute the magnitude of over-education and over-skilling by relying on precise information about a worker's level of attained (*i.e.* completed) education and skills as well as the requirements *to get* (required education) and *to do* (required skills) her/his job. For this purpose, we associate each level of attained education and skills with a given number of equivalent years of education, by relying on the following rule: ISCED_1 is equivalent to primary education, *i.e.* 6 years of education; ISCED_2 relates to lower secondary education, *i.e.* 9 years of education; ISCED_3 is equivalent to upper secondary education, *i.e.* 12 years of education; ISCED_4 to post-secondary non-tertiary education, *i.e.* 14 years of education; ISCED_5 to first stage of tertiary education, *i.e.* 16 years of education; and finally ISCED_6 relates to second stage of tertiary education, *i.e.* 17 years of education. *In fine*, comparing the number of years of attained education (skills) of a worker with the number required *to get* (*to do*) the worker's job gives us the number of years of over-education (over-skilling).

First, in order to investigate the wage effects of over-education, we rely on an extended version of the Mincer wage equation model that is widely used in the existing literature (Chevalier 2003; Frenette 2004; Dolton and Silles 2008; Green and Zhu 2010; Pecoraro 2014; Kracke, Reichelt, and Vicari 2018). Our first equation is formulated as follows:

$$\ln w_i = \alpha_0 + \alpha_1 OverEduY_i + \alpha_2 X_i + \varepsilon_i \tag{1}$$

where:

(a) $OverEduY_i$ is the number of years of over-education of worker i, computed as the difference between the number of years of education or training of worker i and the number of years of education or training needed to get worker i's job if > 0, 0 otherwise.
(b) X_i is a vector of control variables including: the number of years of under-education of worker i, the length of studies (*i.e.* the attained level of education of worker i), the field of education, gender, age, tenure, type of employment contract (respectively, part time or not, and open-ended or not), sector of activity, size of the firm where worker i is employed, and the country in which the firm employing worker i is located.
(c) ε_i is the error term.

Note that, as mentioned by Pecoraro (2014), equation (1) can be seen as a test for human capital theory, according to which the educational requirements for a job do not influence workers' wages as they are solely driven by workers' characteristics. In order to validate human capital theory, over-educated workers should earn the same wage as their peers employed in jobs matching their level of education so that each additional year of over-education should have no impact on wages, that is $\alpha_1 = 0$. However, if human capital theory is rejected, wages are expected to be determined by job requirements, and the returns to over-education should thus be negative and significant, that is $\alpha_1 < 0$. Such models used in the literature mainly show that over-educated workers suffer from a wage penalty in comparison with workers similarly educated but working in jobs matching their level of education. This outcome is more in line with the assignment model than standard human capital theory (Dolton and Vignoles 2000; McGuinness 2006; Leuven and Oosterbeek 2011).

Our second equation investigates the impact of the interaction between over-education and over-skilling on wages, where the measurement of over-education from equation (1) is replaced by a vector of interaction variables including both types of mismatch in the same equation. It is formulated as follows:

$$\ln w_i = \beta_0 + \beta_1 AppOverEdu_i + \beta_2 GenuineOverEdu_i + \beta_3 X_i' + \varepsilon_i \quad (2)$$

where:

(a) $AppOverEdu_i$ represents apparent over-education and is computed as the interaction between two variables: i) a dummy variable taking the value 1 if worker *i* considers herself/himself as properly skilled for her/his job, 0 otherwise; and ii) the number of years of over-education for worker *i* as computed in equation (1). This variable enables us to estimate the wage effect of an additional year of over-education for a worker considered as properly skilled for her/his job, *i.e.* of an additional year of apparent over-education.
(b) $GenuineOverEdu_i$ represents genuine over-education and is computed following the same reasoning, namely the interaction between: i) a dummy variable taking the value 1 if worker *i* considers herself/himself as over-skilled for her/his job, 0 otherwise; and ii) the number of years of over-education for worker *i* as computed in equation (1). This variable allows us estimate the wage effect of an additional year of over-education for a worker considered as over-skilled for her/his job, *i.e.* of an additional year of genuine over-education.
(c) X'_i is a vector of control variables: the number of years of apparent matching of worker *i* (obtained by the interaction of the following two variables: i) a dummy taking the value 1 if worker *i* considers herself/himself as properly educated, 0 otherwise; and ii) the number of years of over-skilling for worker *i* as computed in equation (1)); the length of studies (*i.e.* the attained level of education of worker *i*); field of education, gender; age; tenure; type of employment contract (respectively, part time or not, and open-ended or not); sector of activity; size of the firm where worker *i* is employed; and the country in which the firm employing worker *i* is located.

In this second equation, apparently over-educated and genuinely over-educated workers are compared with each other in situations where they were not mismatched in terms of education and skills.

Equations (1) and (2) have first been estimated using the ordinary least squares method (OLS). The OLS estimator, with standard errors robust to heteroscedasticity and serial correlation, is based on the cross-section variability between workers. Relying on OLS supposes that our mismatch variables are not correlated with unobserved variables (that could also explain wages), such as innate ability or family background. Moreover, it requires to assume that there is no reverse causality, *i.e.* that wages have no effect on mismatch variables. If these assumptions are not satisfied, OLS estimates are biased and inconsistent due to endogeneity. Put differently, statistical inference is not permitted. Some papers address this potential endogeneity bias, but they do so mainly through the use of fixed effects or GMM estimators in the context of panel data (see *e.g.* Carroll and Tani 2013; Kampelmann and Rycx 2012; Kracke and Rodrigues 2020; Marioni 2021; Reis 2017; Tsai 2010). Moreover, these are studies dealing only with the phenomenon of over-education or that of over-skilling. To the best of our knowledge, studies on the interaction between over-education and over-skilling are relatively scarce, and none of them have addressed the issue of potential endogeneity. This is probably due to data availability and to the difficulty of finding a good strategy to tackle the endogeneity of interaction variables between the educational and skills mismatches. For example, Caroleo and Pastore (2018, 1006) mentioned that 'endogeneity […] is not addressed here due to the lack of appropriate longitudinal data'.

Endogeneity is therefore a potential source of bias that deserves to be taken into account in our study. In order to do so, we rely on an instrumental variable estimation strategy, namely the two-stage least squares (2SLS) method. This method requires finding instruments that are both correlated with the (potentially endogenous) variables of interest and uncorrelated with the dependent variable. We rely on the following three instruments: i) willingness to work close to home, ii) preference for leisure activities, and iii) preference for self-learning. On the one hand, these instruments are most likely correlated with our mismatch variables and are therefore relevant. On the other hand, although they may not be strictly exogenous, it is reasonable to assume that their correlation with wages is low enough to control, or at least to reduce, the endogeneity problem. The relevance and (fairly) exogeneous nature of our instruments, discussed in Section 4, is supported by an array of diagnosis tests. On any account, we believe that comparing our OLS and 2SLS estimates is useful to determine (or at least get a better idea of) the extent to which the wage effects of the interaction variables between the educational and skills mismatches are biased due to the potential presence of endogeneity.

3. Materials

In spring 2014, the European Center for the Development of Vocational Training (CEDEFOP) commissioned Ipsos to carry out the first pan-European survey on skills mismatch. The European Skills and Jobs (ESJ) survey was conducted by telephone or online on 48,676 employees aged between 24 and 65 in the 28 European Union member states.

The analyses conducted in this article face some restrictions relating to the data. For instance, workers were asked to report their gross monthly wage during the interviews, but some of them refused to do so. Also, workers who did not report their educational and/or skills level had to be excluded from the dataset. This results in a final sample of 23,123 exploitable observations.

The descriptive statistics for these observations are presented in Table 1. As far as education is concerned, 0.4% of the workers are primary educated, 7.3% lower are secondary educated, 25.2% have attained upper secondary education, 11.1% are post-secondary educated, 48.3% have attained tertiary education at the first level, and 7.7% are tertiary educated at an advanced level. Concerning mismatch variables, 26.4% of the workers consider themselves as over-educated, with mean years of over-education evaluated at 0.9 year, and 28.9% of the workers estimate that they possess some skills in surplus and are thus over-skilled. The interactions between the educational and skills mismatches show that 2.5% of the sample is apparently over-educated (*i.e.* over-educated but properly skilled) and 23.6% of workers are genuinely over-educated (*i.e.* over-educated and over-skilled). Finally, 44.7% of all workers are female, 41.3% have at least 10 years of tenure, 13.5% work in part-time jobs, and the majority of worker are employed in SMEs (*i.e.* firms with less than 250 workers).

4. Results

Table 2 reports the estimation results of equations (1) and (2). Column (1) presents the results of the first equation and focuses on the impact of years of over-education on wages. The adjusted R-squared is equal to 0.45, which suggests that the quality of the fit of our regression model is quite good. What's more, the estimates show that each additional year of over-education leads to a wage penalty of 6.7%[3] in comparison with similarly educated individuals working in a matching situation. This gives support to the idea of a wage penalty associated with over-education but not to simple human capital explanations of wages. Column (2) presents the results of equation (2) and shows that all types of interactions are associated with wage penalties. More precisely, the estimates suggest that an additional year of over-education for apparently over-educated workers is associated with a 4.1%[2] wage penalty, which is lower than the 6.9%[2] penalty for a year of over-education for genuinely over-educated workers.[4] As in Green and Zhu (2010) or Pecoraro

[3]The effect of each variable in percentage is calculated as $e^{\beta_i} - 1$, given the log-linear form of the estimated equation.

[4]The rejection of the null hypothesis that the regression coefficients (when comparing apparent over-education and genuine over-education) are equal gives support to that statement.

Table 1. Descriptive statistics of selected variables.

Variables	Mean	Std. Dev.	Min	Max
Level of attained education (% of workers)				
Primary education	0.39	6.23	0	1
Lower secondary education	7.30	26.01	0	1
Upper secondary education	25.21	43.42	0	1
Post-secondary education	11.10	31.42	0	1
Tertiary education first level	48.26	49.97	0	1
Tertiary education advanced level	7.74	26.72	0	1
Attained education (on average, in years)	14.30	2.37	6	17
Over-education				
Percentage of workers	26.35	44.06	0	1
Years of over-education	0.85	1.68	0	11
Over-skilling				
Percentage of workers	28.92	0.45	0	1
Years of over-skilling	0.93	1.76	0	11
Interaction between the educational and skills mismatches				
Apparently over-educated (% of workers)	2.51	15.64	0	1
Genuinely over-educated (% of workers)	23.64	42.49	0	1
Workers with at least 10 years of tenure (%)	41.27	49.23	0	1
Women (%)	44.70	49.72	0	1
Share of workers < 30 years	43.97	49.64	0	1
Share of workers between 30 and 49 years	43.24	49.54	0	1
Share of workers > 49 years	12.77	33.38	0	1
Part-time (%)	13.50	34.17	0	1
Firm size (% of firms)[a]				
Micro (between 1 and 9 workers)	20.01	40.01	0	1
Small (between 10 and 49 workers)	28.17	44.98	0	1
Medium (between 50 and 249 workers)	25.97	43.85	0	1
Large (>250 workers)	24.18	42.82	0	1
Not mentioned	1.67	12.84	0	1
Sector (% of firms)				
Agriculture, horticulture, forestry or fishing (A)	1.68			
Supply of gas or electricity, mining or quarrying (B + D)	1.91			
Supply, management or treatment of water or steam (E)	0.99			
Manufacturing or engineering (C)	13.63			
Construction or building (F)	5.32			
Retail, sales, shop work or whole sale (G)	9.49			
Accommodation, catering or food services (I)	2.92			
Transportation or storage (H)	5.52			
Information technology or communication services (J)	6.48			
Financial, insurance or real estate services (K + L)	5.44			
Professional, scientific or technical services (M)	7.26			
Administration and support services, including public (N + O)	12.34			
Services relating to education or health (*P*)	18.24			
Cultural industries (arts, entertainment or recreation) (R)	1.96			
Social and personal services (Q)	5.86			
Other	0.96			
Number of observations		23,123		

[a]According to the European standard definitions of medium and small firms.

(2014), the impact of being both over-educated and over-skilled is thus consistently estimated to be more significantly negative than being over-educated but properly skilled. These differences, measured in terms of earning penalties, reinforce the idea that skills heterogeneity matters significantly when analysing the impact of over-education on wages.

When taking endogeneity into account in the relationships through the 2SLS estimator, the first stage regressions (see bottom of the Table 2) first show that the coefficients of the three IVs have the expected sign, when significant. Therefore, being keen to learn alone, favouring a job near home, and favouring leasuring activities all increase the probability of being over-educated or genuinely over-educated. Only the variable of apparent over-education records insignificant coefficients for two IVs ('learning alone' and 'favouring leisure activities'). However, in this case, the coefficient for

Table 2. Educational mismatch, skills mismatch, and wages (OLS estimates, 2014).

	Gross monthly wage (ln)			
	OLS		2SLS	
Dependent variable	Educational mismatch (1)	Educational and skills mismatches (2)	Educational mismatch (1′)	Educational and skills mismatches (2′)
Over-education (in years)	−0.069***		−0.200***	
	(0.004)		(0.048)	
Apparent over-education[a] (in years)		−0.042**		−0.072***
		(0.012)		(0.049)
Genuine over-education[b] (in years)		−0.071***		−0.197***
		(0.004)		(0.049)
Length of studies				
Lower secondary education (dummy)	0.079	0.077	0.084	0.076
	(0.078)	(0.078)	(0.078)	(0.077)
Upper secondary education (dummy)	0.290***	0.286***	0.398***	0.373***
	(0.076)	(0.076)	(0.086)	(0.083)
Post-secondary education (dummy)	0.377***	0.367***	0.568***	0.522***
	(0.079)	(0.079)	(0.106)	(0.099)
Tertiary education first stage (dummy)	0.585***	0.585***	0.762***	0.743***
	(0.078)	(0.078)	(0.101)	(0.099)
Tertiary education advanced stage (dummy)	0.777***	0.778***	0.994***	0.977***
	(0.081)	(0.081)	(0.113)	(0.112)
Other control variables[c]	YES	YES	YES	YES
Sig. model (*p*-value)	0.000	0.000	0.000	0.000
Adjusted R-Squared	0.45	0.45	0.42	0.42
Number of firm-year observations	23,123	23,123	23,123	23,123
2SLS – Diagnosis tests				
Weak identification test:				
Kleibergen-Paap rk Wald F statistic			45.13	46.70
Under-identification test:				
p-value Kleibergen-Paap rk LM statistic			0.000	0.000
Over-identification test:				
p-value of Sargan-Hansen J statistic			0.156	0.103
Endogeneity test:				
p-value associated with Chi-squared statistic			0.007	0.010
2SLS – First stage regressions	Over-education	Apparent over-education	Genuine over-education	
Learn_Alone	0.211***	−0.0001	0.210***	
	(0.022)	(0.008)	(0.021)	
Pref_Home	0.118***	0.019**	0.094***	
	(0.024)	(0.009)	(0.023)	
Pref_Leisure	0.127***	0.005	0.128***	
	(0.023)	(0.008)	(0.022)	

Notes: Robust standard errors are reported between brackets.
***, **, * significant at the 1, 5, and 10% level, respectively.
[a]Apparent over-education means being over-educated (in years) but properly skilled (dummy variable).
[b]Genuine over-education means being over-educated (in years) and over-skilled (dummy variable).
[c]Are included in the vector of control variables, besides the length of study representing the attained level of education: the number of years of under-education of worker i *(1)* or the number of years of apparent matching *(2)*, the experience on the job computed as years of tenure, the fact of being part of the class of workers that is younger than 30 and older than 49 years, respectively. Are also included: the gender (women), working part-time (or not) as well as working under open-end contracts (or not), the study field of the worker (14 dummies), the country where the worker operates (27 dummies), the sectorial affiliation of the firm in which the worker operates (16 dummies), and the size of the firm (*i.e.* the number of workers gathered in 4 dummies).

the ‘favouring a job near home’ IV is also significant and has the expected sign. As for the second stage regressions and the main outcomes (see upper part of Table 2), they confirm, in column (1′), the existence of a wage penalty associated with over-education: one additional year of over-education leads to a 18.1%[2] wage drop in comparison with similarly educated individuals working in matching situations.[5] Including the skills mismatch in the relationship (column (2′)) shows that the wage penalty associated with apparent over-education remains significant with a 6.9%[2] penalty for each additional year of apparent over-education. This reveals, however, that the highest wage penalty is associated with genuine over-education (a 17.9%[2] penalty for a one-year increase in the level of genuine over-education, *i.e.* considering both over-education and over-skilling)[4].

To assess the robustness of the 2SLS approach and the validity of our instruments, we performed an array of diagnosis tests.[6] The results of these tests are reported in Table 2. The first-stage estimates, discussed above, suggest that our IVs are not weak. This is also corroborated by the Kleibergen-Paap rk Wald F statistic for weak identification. This statistic is a Wald F test examining whether the excluded instruments are sufficiently correlated with the endogenous regressor. The null hypothesis is that the instruments are weak. According to the standard 'rule of thumb', weak identification is problematic for F statistics smaller than 10 (as suggested by van Ours and Stoeldraijer (2011)). Our F statistics are higher than 10, for equations (1) and (2) both. Moreover, the Kleibergen-Paap rk LM statistic for under-identification tests whether the equation is identified, *i.e.* whether the excluded instruments are all relevant. The null hypothesis in this test is that the equation is under-identified. The results show that we can reject the null hypothesis that our first-stage equation is under-identified, since the Kleibergen-Paap rk LM statistic is found to be highly significant for both equations. Next, to examine whether our instruments fulfil the exogeneity condition, we computed bivariate correlations between our IVs and our dependent variable. Our findings, available on request, show that all correlation coefficients are very small (between 3% and 7%) and support the assumption that our IVs are fairly exogenous with respect to wages. Concerning the quality of our instruments, the Sargan-Hansen J statistic tests the null hypothesis that the instruments are valid, *i.e.* uncorrelated with the error term. The corresponding *p*-values are equal to 0.156 and 0.103 for equations (1) and (2), respectively, which suggests that our instruments are valid. Finally, the Durbin-Wu-Hausman endogeneity test is based on the difference between two Sargan-Hansen statistics: one for the equation in which mismatch variables are treated as endogenous, and one in which they are treated as exogenous. If the null hypothesis of this test cannot be rejected, then instrumentation is actually not necessary, *i.e.* mismatch variables can actually be considered as exogenous. The *p*-values associated with this endogeneity test are equal to 0.007 and 0.010 for equations (1) and (2), respectively. These results suggest that the null hypothesis of no endogeneity should be rejected. These estimates thus indicate that our main explanatory variables are endogenous and that our instrumentation strategy is relevant.

5. Discussion

Educational mismatch is an important and growing phenomenon in Europe: workers are found to be more and more educated, which results in a risk for some workers to be allocated to jobs that do not match their level of education. A range of authors however argue that education does not reflect the real competencies of workers on the job (Chevalier and Lindley 2009; Verhaest and Omey 2009). This has led researchers to investigate the skills mismatch phenomenon, whereby workers are considered as possessing (or not) the skills needed to perform their job (Mavromaras, McGuinness, and Wooden 2007; OECD 2011; Caroleo and Pastore 2018).

Relying on the European Skills and Jobs (ESJ) survey, *i.e.* the first pan-European survey on skills mismatch, this paper investigates the wage effects of new, more precise measurements of mismatch. Whereas most of the current investigations rely solely on workers' educational background, this

[5]Such large effect may be somewhat suprising. However, it should be noted that a one-year increase in the level of over-education implies more than doubling the current situation (the descriptive statistics show 0.85 mean year of over-education). The leverage is therefore all the more important.

[6]It should be noted that finding good instruments, *i.e.* variables that are both relevant and exogenous, remains a difficult task. Therefore, although the diagnosis tests support the validity of our instruments, 2SLS results should be interpreted with caution.

paper challenges education and skills by interacting over-education with over-skilling to distinguish two specific situations: (i) apparent over-education (*i.e*. when a worker is over-educated but properly skilled) and (iii) genuine over-education (*i.e.* when a worker is both over-educated and over-skilled). Taking into account both the skills mismatch and the more classical educational mismatch allows us to account for workers' heterogeneity in their skills and makes this paper one of the first, among those using the ESJ survey, to evaluate the combined effect of over-education and over-skilling on wages from a European perspective. We also deepen the analysis by computing the magnitude of the wage effects of the two mismatches in terms of additional years of over-education, apparent over-education, and genuine over-education. Finally, we are among the first to test for the existence of a potential endogeneity bias associated with education and skills mismatch variables.

Our estimates first show that over-educated workers, *ceteris paribus*, earn less than their opposite numbers employed in jobs matching their level of education. More precisely, we find, on the basis of 2SLS estimates, that each additional year of over-education results in a wage penalty of around 18% across EU countries. Yet, this penalty is found to vary substantially depending on whether over-education coincides with over-skilling. One year of over-education for workers that are properly skilled (*i.e.* apparently over-educated) leads to a non-negligible wage penalty of almost 7%, whereas the penalty for genuine over-education is more than two and a half times greater. These estimates, controlling for a large number of covariates and addressing the endogeneity issue, are of concern in several respects. First, our descriptive statistics show that genuinely over-educated workers represent a large group of mismatched workers in the EU. This category represents over 20% of the workforce, on average. Obviously, this is quite costly from a worker's perspective. But the cost of mismatch is also likely to be huge for the economy overall. Recent papers have indeed shown that the wage effects of the educational and skills mismatches reflect (to a large extent) differences in productivity among workers. Put differently, they indicate that over-educated/over-skilled workers would create much more value added if they were employed in jobs matching their education/skills. Along those lines, our estimates suggest that a better allocation of resources in the economy (*i.e.* notably through a decrease in the incidence of genuinely over-educated workers) could improve overall productivity significantly. Improving the quality of job-workers matches in EU countries thus appears to be a key challenge not only from a worker's perspective but also for the economy at large.

Disclosure statement

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

ORCID

L. Cultrera http://orcid.org/0000-0003-4400-4524
F. Rycx http://orcid.org/0000-0003-0964-0939
G. Vermeylen http://orcid.org/0000-0001-9641-060X

References

Allen, J., and R. van der Velden. 2001. "Educational Mismatches Versus Skill Mismatches: Effects on Wages, Job Satisfaction, and On-the-Job Search." *Oxford Economic Papers* 3: 434–452. doi:10.1093/oep/53.3.434.

Barro, R. J., and J. W. Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics* 104: 184–198. doi:10.1016/j.jdeveco.2012.10.001

Bauer, T. K. 2002. "Educational Mismatch and Wages: A Panel Analysis." *Economics of Education Review* 21: 221–229. doi:10.1016/j.jdeveco.2012.10.001.

Bender, K. A., and J. S. Heywood. 2011. "Educational Mismatch and the Careers of Scientists." *Education Economics* 19: 253–274. doi:10.1080/09645292.2011.577555.

Caroleo, F. E., and F. Pastore. 2018. "Overeducation at a Glance: Determinants and Wage Effects of the Educational Mismatch, Looking at the AlmaLaurea Data." *Social Indicators Research* 137: 999–1032. doi:10.1007/s11205-017-1641-1.

Carroll, D., and M. Tani. 2013. "Over-Education of Recent Higher Education Graduates: New Australian Panel Evidence." *Economics of Education Review* 32: 207–218. doi:10.1016/j.econedurev.2012.10.002.

Chevalier, A. 2003. "Measuring Over-education." *Economica* 70: 509–531. doi:10.1111/1468-0335.t01-1-00296.

Chevalier, A., and J. Lindley. 2009. "Overeducation and the Skills of UK Graduates." *Journal of Royal Statistical Society* 172: 307–337. doi:10.1111/j.1467-985X.2008.00578.x.

Cohn, E., and S. P. Khan. 1995. "The Wage Effects of Overschooling Revisited." *Labour Economics* 2: 67–76. doi:10.1016/0927-5371(95)80008-L.

Dolton, P. J., and M. A. Silles. 2008. "The Effects of Over-education on Earnings in the Graduate Labour Market." *Economics of Education Review* 27: 125–139. doi:10.1016/j.econedurev.2006.08.008.

Dolton, P., and A. Vignoles. 2000. "The Incidence and Effects of Overeducation in the U.K. Graduate Labour Market." *Economics of Education Review* 19: 179–198. doi:10.1016/S0272-7757(97)00036-8.

Duncan, G. J., and S. D. Hoffman. 1981. "The Incidence and Wage Effects of Overeducation." *Economics of Education Review* 1: 75–86. doi:10.1016/0272-7757(81)90028-5.

European Commission. 2017. *Employment and Social Developments in Europe 2017*. Luxembourg: Publications Office of the European Union.

European Commission. 2019. *Education and Training – Monitor 2019*. Luxembourg: Publications Office of the European Union.

Freeman, R. B. 1976. *The Overeducated American*. New York: Academic Press.

Frenette, M. 2004. "The Overqualified Canadian Graduate: The Role of the Academic Program in the Incidence, Persistence, and Economic Returns to Overqualification." *Economics of Education Review* 23: 29–45. doi:10.1016/S0272-7757(03)00043-8.

Green, F. 2007. *Demanding Work: The Paradox of Job Quality in the Affluent Economy*. Princeton, NJ: Princeton University Press.

Green, F., S. McIntosh, and A. Vignoles. 1999. *'Overeducation' and Skills: Clarifying the Concepts*. Discussion Paper 435. London: Centre for Economic Performance.

Green, F., and Y. Zhu. 2010. "Overqualification, Job Dissatisfaction, and Increasing Dispersion in the Returns to Graduate Education." *Oxford Economic Papers* 62: 740–763. doi:10.1093/oep/gpq002.

Groot, W. 1993. "Overeducation and the Returns to Enterprise-related Schooling." *Economics of Education Review* 12: 299–309. doi:10.1016/0272-7757(93)90064-N

Groot, W., and H. Maassen van den Brink. 2000. "Overeducation in the Labor Market: A Meta-Analysis." *Economics of Education Review* 19: 149–158. doi: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: 131–147. doi:10.1016/S0272-7757(99)00050-3

Kampelmann, S., and F. Rycx. 2012. "The Impact of Educational Mismatch on Firm Productivity: Evidence from Linked Panel Data." *Economics of Education Review* 31: 918–931. doi:10.1016/j.econedurev.2012.07.003.

Kiker, B. F., M. C. Santos, and M. Mendes de Oliveira. 1997. "Overeducation and Undereducation: Evidence for Portugal." *Economics of Education Review* 16: 111–125. doi:10.1016/S0272-7757(96)00040-4

Korpi, T., and M. Tahlin. 2009. "Educational mismatch, wages and wage growth: overeducation in Sweden, 1974-2000." *Labour Economics* 16: 183–193. doi:10.1016/j.labeco.2008.08.004.

Kracke, N., M. Reichelt, and B. Vicari. 2018. "Wage Losses due to Overqualification: The Role of Formal Degrees and Occupational Skills." *Social Indicators Research* 139: 1085–1108. doi:10.1007/s11205-017-1744-8.

Kracke, N., and M. Rodrigues. 2020. "A Task-Based Indicator for Labour Market Mismatch." *Social Indicators Research* 149: 399–421. doi:10.1007/s11205-019-02261-2

Leuven, E., and H. Oosterbeek. 2011. "Overeducation and Mismatch in the Labor Market." Chap. 3 in *Handbook of The Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 283–326. Amsterdam: Elsevier. doi:10.1016/B978-0-444-53444-6.00003-1.

Lindley, J., and S. McIntosh. 2010. "Is the Over-Education Wage Penalty Permanent?" Sheffield Economic Research Paper Series No. 2010004.

Mahy, B., F. Rycx, and G. Vermeylen. 2015. "Educational Mismatch and Firm Productivity: Do Skills, Technology and Uncertainty Matter?" *De Economist* 163: 233–262. doi:10.1007/s10645-015-9251-2

Marioni, L. D. S. 2021. "Overeducation in the Labour Market: Evidence from Brazil." *Education Economics* 29: 53–72. doi:10.1080/09645292.2020.1832201.

Mateos-Romero, L., and M. M. Salinas-Joménez. 2018. "Labor Mismatches: Effects on Wages and on Job Satisfaction in 17 OECD Countries." *Social Indicators Research* 140: 369–391. doi:10.1007/s11205-017-1830-y.

Mavromaras, K., S. McGuinness, and Y. K. Fok. 2009. "Assessing the Incidence and Wage Effects of Overskilling in the Australian Labour Market." *Economic Record* 85: 60–72. doi:10.1111/j.1475-4932.2008.00529.x.

Mavromaras, K., S. McGuinness, N. C. O'Leary, P. J. Sloane, and Z. Wei. 2013. "Job Mismatches and Labour Market Outcomes: Panel Evidence on University Graduates." *Economic Record* 89: 382–395. doi:10.1111/1475-4932.12054.

Mavromaras, K., S. McGuinness, and M. Wooden. 2007. "Overskilling in the Australian Labour Market." *Autralian Economic Review* 40: 307–312. doi:10.1111/j.1467-8462.2007.00468.x.

McGuinness, S. 2006. "Overeducation in the Labour Market." *Journal of Economic Surveys* 20: 387–418. doi:10.1111/j.0950-0804.2006.00284.x.

McGuinness, S., and D. Byrne. 2014. "Examining the Relationships between Labour Market Mismatches, Earnings and Job Satisfaction among Immigrant Graduates in Europe." IZA Discussion Paper No. 8440.

McGuinness, S., K. Pouliakas, and P. Redmond. 2018. "Skills Mismatch: Concepts, Measurement and Policy Approaches." *Journal of Economic Surveys* 32: 985–1015. doi:10.1111/joes.12254.

OECD. 2011. *OECD Employment Outlook 2011 – Chapter 4: Right for the Job – Over-Qualified of Under-Skilled?* Paris: OECD Publishing.

Pecoraro, M. 2014. "Is There Still a Wage Penalty fo Being Overeducated But Well-Matched in Skills? A Panel Data Analysis of Swiss Graduate Cohort." *Labour (committee. on Canadian Labour History)* 28: 309–337. doi:10.1111/labr.12031.

Pecoraro, M. 2016. "The Incidence and Wage Effects of Overeducation using the Vertical and Horizontal Mismatch in Skills: Evidence from Switzerland." *International Journal of Manpower* 37: 536–555. doi:10.1108/IJM-10-2014-0207.

Pellizzari, M., and A. Fichen. 2013. "A New Measure of Skills Mismatch: Theory and Evidence from the Survey of Adult Skills (PIAAC)" OECD Social, Employment and Migration Working Papers No 153.

Quintini, G. 2011. "Over-Qualified or Under-Skilled: A Review of Existing Literature." OECD Social, Employment and Migration Working Papers No 121.

Reis, M. C. 2017. "Educational Mismatch and Labor Earnings in Brazil." *International Journal of Manpower* 38: 180–197. doi:10.1108/IJM-02-2016-0030.

Rumberger, R. 1987. "The Impact of Surplus Schooling on Productivity and Earnings." *Journal of Human Resources* 22: 24–50. doi:10.2307/145865

Sicherman, N. 1991. "Overeducation in the Labor Market." *Journal of Labor Economics* 9: 101–122. doi:10.1086/298261

Sloane, P. J., H. Battu, and C. Belfield. 1999. "Overeducation Among Graduates: A Cohort View." *Education Economics* 7: 21–38. doi:10.1080/09645299900000002.

Tsai, Y. 2010. "Returns to Overeducation: A Longitudinal Analysis of the US Labor Market." *Economics of Education Review* 29: 606–617. doi:10.1016/j.econedurev.2010.01.001.

van Ours, J., and L. Stoeldraijer. 2011. "Age, Wage and Productivity in Dutch Manufacturing." *De Economist* 159: 113–137. doi:10.1007/S10645-011-9159-4.

Verdugo, R., and N. Verdugo. 1989. "The Impact of Surplus Schooling on Earnings: Some Additional Findings." *Journal of Human Resources* 24: 629–643. doi:10.2307/145998

Verdugo, R., and N. Verdugo. 1992. "Surplus Schooling and Earnings: Reply to Cohn and to Gill and Solberg." *Journal of Human Resources* 27: 690–695. doi:10.2307/146083

Verhaest, D., and E. Omey. 2009. "Objective Over-education and Worker well-being: A Shadow Price Approach." *Journal of Economic Psychology* 30: 469–481. doi:10.1016/j.joep.2008.06.003