

"This is an Author's Accepted Manuscript of an article published
in , Oxford Review of Education, Vol. 46, Issue 1, 2020, pp. 10-29
[copyright Taylor & Francis], available online at:
<https://doi.org/10.1080/03054985.2019.1687433>

Educational expansion and overeducation of young graduates: A comparative analysis of 30 European countries

Abstract

This study uses quarterly EU Labour Force Survey data for 30 countries over the period 2000 to 2016 to examine the relationship between changes in the composition of educational attainment and overeducation rates among new labour market entrants holding post-secondary and tertiary qualifications. We find that tertiary education expanded rapidly across our sample, while the proportion of young people with lower levels of education fell gradually throughout the period. Despite the significant increases in the percentage of young people educated to tertiary level, overeducation among new tertiary graduates fell. The descriptive evidence also suggests that some of the greatest declines in overeducation of young tertiary graduates occurred in the countries experiencing the most significant expansion in tertiary education. Overeducation rates among young graduates with upper-secondary and post-secondary (non-tertiary) education were lower in magnitude than overeducation of tertiary graduates and declined slightly over the period. Our GMM results confirm the negative relationship between educational expansion and overeducation for both tertiary and post-secondary graduates, and reveal a number of other factors potentially explaining cross-country variation in youth overeducation rates.

Key words; overeducation; educational expansion; gmm

JEL codes: I28; C23

1.0 Introduction

It is often assumed that overeducation is a recent phenomenon that has become more prominent due to increasing higher education participation rates in most developed countries in recent decades. However, McGuinness et al. (2018a) show that overeducation rates have remained relatively unchanged over time in many EU countries and may actually be declining in others. They report that the incidence of overeducation in the EU, averaged over all countries and education levels, has remained stable at approximately 18 per cent from 2003 to 2013. There is also a perception that overeducation predominantly affects tertiary graduates and the existing literature tends to be focused in this direction (see e.g., Chevalier and Lindley, 2009; Croce and Ghignoni, 2012; Baert et al., 2013; Li and Miller, 2015; Carroll and Tani, 2015). The fact that overeducation can also occur at lower levels of educational attainment has been largely overlooked in research to date. However, in the policy analysis literature several authors have noted that overeducation rates in some EU countries (e.g. Spain, Portugal, Italy, Romania) have been disproportionately concentrated among mostly medium-educated workers, possibly related to the strength of their vocational education and training (VET) systems (Pouliakas, 2012; Lessaer et al., 2015).

This paper uses unique data from 30 EU countries to shed light on a number of previously under-researched issues regarding the evolution and drivers of overeducation rates among young people at various levels of educational attainment. Our data and adopted empirical approach allows us to examine these issues within, and across, European countries over time. To date, the examination of such issues has been limited by the absence of time-series data on international overeducation rates. To overcome the data deficit, we build on the work of McGuinness et al. (2018a) by using the micro level information contained within quarterly EU Labour Force Surveys, to construct an overeducation series of young graduates at various levels of educational attainment for 27 EU countries plus Iceland, Switzerland and Norway over time.¹ Throughout the text, we will refer to this group of 30 countries as EU+ countries. Using this database, we study a number of issues that have yet to be examined in the international literature. We make three main contributions. Firstly, we examine the extent to which the educational profile of young people, many of which are new labour market entrants, has been altering over time within the EU+ countries. Secondly, we trace the evolution of overeducation

¹ We have all EU member states apart from Malta due to data limitations.

rates of young graduates by educational attainment level across European countries over time. In addition to high-educated (tertiary) graduates, we focus also on medium-qualified workers, educated to upper-secondary and post-secondary (non-tertiary) level. Thirdly, we investigate the relationship between overeducation rates at different educational attainment levels and the composition of labour supply within the EU+ countries. In particular, we examine the relationship between the provision of post-secondary educational pathways, which often tend to reflect the strength and quality of VET systems in some European countries, and the risk of youth graduate overeducation, controlling for a range of other potentially confounding factors.

2.0 Previous Literature

Overeducation refers to the phenomenon whereby workers are in jobs for which they hold an educational qualification at a level that exceeds the educational requirements of the job. There exists a substantial literature on overeducation that has been summarised in a number of reviews including McGuinness (2006), Quintini (2011), Pouliakas (2012) and McGuinness et al. (2018b). These reviews give a thorough summary of the existing international evidence on overeducation, allowing us to ascertain a number of stylised facts relating to issues such as the impact of overeducation on earnings, job satisfaction and job mobility.² The most recent review by McGuinness et al. (2018b) examines 98 studies of overeducation, covering approximately 40 mainly high-income countries. They report an average overeducation rate, across all studies, of 24 per cent. McGuinness et al. (2018b) also document the evidence on the overeducation pay penalty, whereby overeducated workers may have lower earnings than workers with the same level of education in matched occupations. They collect 74 separate estimates of the overeducation pay penalty and from these, calculate an average pay penalty of 13.6 percent.

The majority of overeducation studies surveyed by McGuinness et al. (2018b) are cross-sectional in nature, examining the incidence, impact and determinants of overeducation within particular countries. The reason behind the dominance of cross-sectional studies relates to a lack of datasets that allow for the detailed measurement of overeducation across countries over-time. As a result, there is little research aimed at examining, or explaining, international differences in overeducation. However, there exists some recent literature that attempts to fill this gap by taking a cross-country approach.

² These reviews also discuss the various measurement approaches to overeducation as well as the role of unobserved heterogeneity, which may potentially bias some estimates of overeducation.

Verhaest and Van der Velden (2013) use graduate data for 13 European countries and Japan, and find that graduate overeducation is lower in countries with education systems focused on providing students with occupational specific, as opposed to general, skills. They also find that country-level differences can be partly explained by business cycle variations, with overeducation more common during recessions, as well as by the over-supply of highly skilled labour. Croce and Ghignoni (2012), using data for 26 countries from the European Community Household Panel Survey (ECHP) for the period 1998 to 2003, also show that business cycle effects play a role in explaining cross-country variation in overeducation rates.

McGuinness et al. (2018a), using a panel data set for 27 EU countries constructed using quarterly Labour Force Surveys, find that overeducation is lower in countries with higher female participation rates. As such, it is argued that equality legislation and affordable childcare that prevent females from having to occupationally downgrade are important policy levers for preventing overeducation. McGuinness et al. (2018a) also find that the composition of employment plays an important role, with overeducation rates lower in countries with a strong manufacturing sector. There was also regional variation, with lower overeducation rates observed in Eastern and Peripheral European countries that had more developed VET pathways. Davia et al. (2017) use EU-SILC data to also examine regional variations in overeducation rates and, consistent with McGuinness et al. (2018a), find that they tend to be lower for females in regions with strong levels of employment protection legislation. Davia et al. (2017) also find that overeducation is higher in regions with more migrant labour.

Ordine and Rose (2017) show descriptively that there is no strong relationship between country level overeducation rates and the share of individuals with tertiary education. Ordine and Rose (2017) also develop a theoretical model which shows that, in regions with low levels of investment in higher education, reducing incentives to university participation can actually exacerbate the problem of overeducation. More recently, Budria and Moro-Egido (2018) using cross-sectional data from the Cedefop European Skills and Jobs Survey report that overeducation is lower in countries with higher shares of part-time workers, a larger public sector, a higher ratio of high to low skilled workers, a lower labour force participation rate and a lower share of temporary contract workers.

In summary, while the literature on the determinants of cross-country variation in overeducation is somewhat limited, some common findings are beginning to emerge from existing studies. Differences in overeducation rates appear to be related to business cycle

effects, the composition of labour demand, degree of labour market protection and flexibility (dependent on the state of employment protection legislation) and the educational composition of labour supply within a particular country. It is this last relationship, focusing on the role of educational attainment and its expansion as a determinant of youth overeducation rates, that we explore in more depth within the current study.

3.0 Data and methods

There are three ways to measure overeducation: the subjective method, the job evaluation method and the empirical method. The subjective approach is based on survey questions, where respondents are asked to compare their highest level of education with qualifications required to either do, or get, their current job. The job evaluation method compares individual educational attainment with occupational job entry requirements, that are based on the assessments of professional job analysts in the context of drafting occupational dictionaries (such as DOT or O*NET in the United States, SOC in the UK or, more recently, ESCO, in the EU). The empirical method, also referred to as the statistical method, compares an individual's highest educational attainment with the mean, or the modal, level of education within the worker's (usually broadly defined group of) occupation.³ The choice of measurement approach is typically constrained by data availability.

Due to an absence of subjective or occupational dictionary-based measures for EU countries over time, we apply the empirical method to the quarterly waves of the European Labour Force survey (EU-LFS) to generate country specific estimates of overeducation rates among full-time employees, aged 15 to 29, for the period 2000 to 2016.⁴ McGuinness et al. (2017) show that this approach, when applied to EU-LFS data, yields very similar country level estimates compared to PIAAC data using the same approach. Regarding the comparability of the three approaches, in a recent comprehensive review of the overeducation literature, McGuinness, et al. (2018) show that, on average, all three approaches produce very similar estimates. The average overeducation rate collected from a large number of studies across 39 countries by McGuinness et al. (2018) yield the following estimates of overeducation: subjective measure (22%), realised matches (26%), occupational dictionary measure (26%). Nevertheless, it

³ A detailed discussion on the different measurement approaches can be found in McGuinness et al. (2018b), including a discussion of their relative advantages and disadvantages.

⁴ The 2016 survey is the latest dataset available to us.

should be acknowledged that there is not a perfect correlation between measurement approaches and the method selected by researchers is generally dependant on the nature of the data available (see McGuinness (2006), Quintini (2001) and McGuinness et al. (2018) for relevant reviews).

This will represent the key dependent variable in the study. It is the case that some of the individuals within our sample may still be moving through the educational system on a part-time basis or, alternatively, will choose to return to full-time education at some point in the future. Given this, the overeducation status may well change as some point in the future, however, this is potentially the case for workers of all ages. We are of the view that restricting our sample to full-time employees limits this possibility that individuals will still be participating in education and, furthermore, that restricting our sample to somewhat older workers, for aged example 25 to 34, will not accurately reflect the incidence of youth overeducation rates. This is demonstrated by our descriptive data which shows that during the period, the vast majority of young people completed education to ISCED 4 or below, which will occur before the age of 20 in most cases (Figure 2).

The EU-LFS is a large household survey providing quarterly data on workers aged 15 and over as well as on persons outside the labour force. It is conducted in the 28 Member States of the European Union, 2 candidate countries and 3 countries of the European Free Trade Association (EFTA). The Labour Force Surveys are conducted by the national statistical institutes across Europe and are centrally processed by Eurostat. For each country, in each quarter, overeducation is defined as the proportion of employees in employment whose highest education qualification level (as described according to the international standard classification of education - ISCED) lies one level or more above the occupational mode. The occupational modal level of education is the most common qualification possessed by workers in each two-digit occupation group. Overeducation is calculated within two-digit occupational codes and using the five ISCED categories of ≤ 2 , 3, 4, 5B and 5A+6. Thus, if the modal level of schooling in a particular two-digit occupation was measured at ISCED 3, then all individuals educated to ISCED levels 4 and above would be deemed to be overeducated, using our empirical approach.

It could be argued that rapid increases in educational attainment may increase the mode in a given occupation. This could lead to an endogenous, statistical, impact on overeducation rates, whereby overeducation falls because the modal occupation level increases, as opposed to any substantive change to the skill content or educational requirements of the occupation itself. However, McGuinness et al. (2018a), using the same estimation method and data, have shown

that occupation modal levels of education in their sample of advanced economies are relatively constant over time, despite the rapid expansion of educational attainment.

While there has been some variation in the literature concerning the measure of central tendency adopted as a cut-off point, the use of the occupational mode is preferred when education is measured categorically, as in this case a mean value is somewhat difficult to interpret. When education is measured in terms of years of schooling, the use of the occupational mean is more reasonable. While the use of the mode as a cut-off point is not without its problems, it seems the most rationale choice here given the data. As we are interested in distinguishing overeducation by level of educational attainment, for each quarter and for each country, we extract the respective overeducation rate for tertiary graduates (educated to ISCED levels 5A, 5B & 6), upper-secondary and post-secondary (non-tertiary) graduates (ISCED Levels 3 & 4) and individuals with below upper secondary education (ISCED levels 1 and 2). Ideally, we would like to distinguish post-secondary graduates in the data according to their educational orientation (i.e. vocational and general educational attainment), however it is not possible to make this distinction before the 2014 wave due to different country level approaches to mapping the rich diversity of post-secondary (vocational) qualifications into the ISCED framework. For instance, in many countries, such as Ireland where the main form of vocational education is the Post Leaving Certificate, this is classified as an ISCED level 4 qualification. In the UK, most post-secondary vocational programmes appear to be classified under ISCED 3, with very little attainment captured at ISCED level 4. Consequently, we combine ISCED levels 3 & 4 to form an upper and post-secondary, pre-tertiary, category.

In addition to providing descriptive evidence on the evolution of overeducation of young employees with varying levels of schooling over time at various ISCED levels and how these have developed alongside changes in patterns of educational attainment, we also use econometric techniques to examine the determinants of cross-country variations in overeducation. Given that we have panel data, with repeated observations at the country level over time, the data structure allows us to estimate causal relationships using a generalised methods of moments (GMM) estimator (Arellano and Bond (1991)) that controls for both reverse causality and unobserved heterogeneity. The model can be written as:

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \beta_j \Delta x_{ijt} + \Delta \varepsilon_{it} \quad (1)$$

Where y_{it} is the education specific rate of overeducation observed for country i in time t , and y_{it-1} is its one period lagged value. X_{it} represents a number of j independent variables with β_j the associated coefficients and ε_{it} is the error term. Equation (1) represents a first-difference model which removes time-invariant fixed effects. However, the error term may still be correlated with the lagged first-difference of the dependent variable. To overcome this, deeper lags of the dependent variable are used as instruments for differenced lags of the dependent variable. In this regard, the deeper lags will be correlated with the lagged dependent variable but not the error term. Progressively deeper lags are applied until exogeneity is achieved.

Alternative estimators, such as a fixed effect estimator or a random effects with a Mundlak correction, can deal with unobserved heterogeneity but do not deal with the possibility that some of the right-hand variables may be prone to reverse causality, particularly in the case of controls related to labour market participation. The GMM estimator is preferred to other models as it helps to control for both time invariant unobserved heterogeneity and reverse causality. This is achieved by first differencing the model, which removes the fixed effect, and using lags to instrument the explanatory variables in the model. Furthermore, the GMM approach also has the advantage of allowing us to estimate the dynamic persistence of youth overeducation

In terms of explanatory variables, we include in our models control variables that reflect the level, composition and interaction of both labour demand and supply, as well as indicators of the degree of labour market flexibility. Most of the variables are derived from the LFS micro data, such as supply indicators measuring the educational composition of labour supply and labour force participation and the share of female and migrant workers; demand indicators derived from the data include the share of employment across various key sectors and the ratio of high to low skilled occupational employment⁵; the degree of labour market flexibility is captured by the shares of part-time employees. We also include Research and Development expenditure in the models to proxy for the degree of technological sophistication within the labour market. Finally, in keeping with the existing literature, business cycle effects are also controlled for by the inclusion of measures of GDP per capita and the unemployment rate. The per capita GDP data is sourced externally from published annual figures and interpolated to fit the quarterly structure of our data.

⁵ This is defined as the share of workers in SOC 2 and 3 to the share of workers in SOC 7, 8 and 9.

It would have been preferable to include additional policy variables into our analysis, such as measures of indicators of Active Labour Market Policy (ALMP) and Employment Protection Legislation. However, these types of measures do not tend to be measured consistently across countries and, in instances where they are available, the indicators are not collected quarterly to match our data. For example, when annual measures of Employment Protection Legislation (EPL) are interpolated into quarterly datasets, they tend to perform poorly in panel models, due to the low levels of variation in the measures over time, meaning that any impacts are likely to be absorbed into country level fixed effects (McGuinness et. al, 2016).

4.0 Results

4.1. Descriptive statistics

Figure 1 plots the composition of educational attainment among young people aged 15 to 29 from 2000 to 2016 for the 30 European countries in our sample. We use this age range as it is consistent with the ages we use to calculate graduate overeducation rates. However, in Figure 2 we show educational attainment for those aged 24 to 29. Both figures show the same pattern but Figure 2 provides a clearer picture of how educational attainment patterns are changing over time as most young people in this range will have completed their full-time education. The patterns are quite strong; while we see gradual declines in primary, secondary and post-secondary educational attainment, the share of young people holding tertiary attainment qualifications increased rapidly over the period from 11 to 18 percent. While post-secondary education remained the dominant form of educational attainment, its share fell from 47 percent in 2000 to 44 percent in 2016.

Figure 3 plots the average overeducation rates of young full-time employees, aged 15 to 29, by ISCED level, between the year 2000 and 2016. It is important to note that we use the full sample of all full-time employees aged 15 to 65 when calculating the ISCED modal level. Then we calculate overeducation rates by ISCED level for those aged 15 to 29. For instance, the post-secondary category will include individuals aged 15 to 29 with ISCED level 4 attainment who are overeducated because they are employed in an occupation with a modal education level of ISCED 3 or below. As expected, overeducation is highest among young workers educated to tertiary level and lowest for those employees educated to primary or less⁶. However, despite

⁶ Overeducation can only occur for this group in the case of individuals educated to ISCED level 2 who are employed in occupations with a modal education level of ISCED 1.

the rapid increase in tertiary education attainment, overeducation among young employees who have graduated at tertiary level decreased by a third over the period from approximately 60 percent in 2000 to around 40 percent in 2016, providing an initial indication that EU labour markets have been able to absorb recent increases in university graduate labour supply. Overeducation rates among post-secondary qualified employees fell from 20 to 15 percent over the period while overeducation also declined from 13 to just over 2 percent for young workers holding ISCED 2 qualifications. There was no evidence that overeducation levels among any of the groups altered in any significant manner during the period of the global recession that peaked around 2008 and 2009, consistent with the analysis of Pouliakas (2012).

< **Insert Figure 1** >

< **Insert Figure 2** >

The descriptive, aggregate level, analysis discussed above, may mask unobserved heterogeneity at the country level. Tables 1 and 2 hence provide information on changes in both educational attainment and overeducation among young people aged 15 to 29 in the year 2000 and 2016 for each country. Given the declining share of young people qualified to primary level and below and the low rates of overeducation among this group, our analyses will focus on the tertiary and upper/post-secondary graduate labour markets. Table 1 provides, for the year 2000 and 2016, the shares of individuals aged 15 to 29 educated to upper-/post-secondary level and the overeducation rates for full-time employees in this same age and educational attainment group. The shares of young medium-qualified people in 2000 varied greatly from just over 22 percent in Portugal and 32 percent in Spain to over 65 percent in Czechia and Slovak Republic. Above average rates of upper/post-secondary educational attainment appear to be higher in Eastern European countries but are not exclusive to them, with countries such as Germany and Switzerland, with strong initial VET systems, also producing large proportions of young people educated to post-secondary level.

With respect to overeducation among upper/post-secondary qualified workers, this was approximately zero in 12 countries in 2000, which demonstrates that ISCED level 3 effectively represents a qualifications floor in many countries, with very small proportions of full-time employees qualified to ISCED levels 1 and 2. Countries exhibiting zero, or very close to zero, upper/post-secondary overeducation rates include Denmark, Czech Republic, Finland, France, Netherlands, Slovenia and the UK. Overeducation among young full-time employees with upper/post-secondary education was highest in Spain in 2000 and was also high in Portugal,

Ireland, Italy, Greece and Iceland. By 2016 upper/post-secondary educational attainment had fallen in most, but not all, countries. Between 2000 and 2016 the share of upper/post-secondary education increased in Portugal, Spain, Iceland, France, the UK and Ireland and remained somewhat static in Finland and Slovenia. The upper/post-secondary educational share of young people fell by 10 percentage points and over in Czechia, Denmark, Luxembourg, Norway and Slovak Republic. Consistent with Figure 2, overeducation rates of upper/post-secondary qualified workers fell in most countries and were particularly large in Cyprus, Denmark, Iceland, Netherlands, Ireland, Slovenia and the UK. However, youth overeducation also rose significantly in a number of countries over the period under examination, most notably in Hungary, Estonia, Germany, Lithuania and Sweden. In neither period is there any obvious relationship between the share of young individuals with upper/post-secondary education attainment and the incidence of overeducation among young workers educated to upper/post-secondary level.

Table 2 repeats the same exercise for ISCED level 5 graduates and above. In 2000 the share of young persons aged 15 to 29 educated to tertiary level ranged from around 5 percent in Austria, Czechia, Hungary, Italy, Latvia, Poland, Portugal, Romania and Slovak Republic to over 19 percent in France, Norway, Spain, Sweden and the UK. There is no clear pattern between the overeducation rates of young workers in 2000 and the incidence of ISCED 5 level education attainment. Countries exhibiting the highest overeducation rates among young tertiary graduates, such as Austria, Spain, Italy, the UK and Ireland, varied by both geographic location and the incidence of ISCED 5 attainment. Between 2000 and 2016 the share of young people educated to tertiary level increased in every country with the exception of Finland, Lithuania and Norway where it remained roughly static. Countries such as Austria, Cyprus, Czechia, Latvia, Poland and Slovak Republic experienced increases in excess of 10 percentage points. Despite such large increases in educational attainment, overeducation fell in most countries between 2000 and 2016, with some of the largest falls occurring in countries such as Belgium, Cyprus, Denmark, Estonia, Germany, Ireland and the Netherlands, Spain and Switzerland. However, overeducation rose in other countries, including Hungary, Iceland, Poland, Portugal, Romania and Slovak Republic.

4.2. *Econometric estimates*

The econometric estimates from the estimation of the GMM model are provided in Table 3.

4.2.1. *Young tertiary level graduates*

Dealing firstly with the results for ISCED level 5 graduates, reported in column 1 of Table 3, the lagged dependent variable is positive and highly significant, which is consistent with previous studies showing the dynamic nature of overeducation (see, e.g., Mavromaras and McGuinness, 2012). A striking aspect of the model is the inverse relationship between youth overeducation of university graduates and the share of full time employees who are aged 15 to 29 with third level qualifications. This is consistent with our descriptive evidence and seems to suggest that European labour markets have been generally able to absorb the rapid increases in the supply of highly qualified young people in recent decades. This is an important finding, as it goes against the widely held assumption that overeducation is largely a consequence of rapidly expanding higher educational supply outstripping the demand for graduate labour. Conversely, an increase in the share of young employees with post-secondary qualifications is associated with increased university graduate overeducation. What this implies is that an increase in the relative number of jobs requiring skills and qualifications at ISCED 3 and 4 levels may be conducive to more university graduates occupying these types of jobs, hence resulting in greater tertiary graduate overeducation.

< Insert Table 3 >

In terms of the other explanatory variables, we find that levels of labour market flexibility, proxied by the share of temporary employment, and openness (share of migrants) are significant or marginally insignificant, but that these tend to work in opposite directions. While not statistically significant at conventional levels, youth graduate overeducation is positively related to the share of temporary employment in the labour market.⁷ This may reflect that for some workers a temporary contract can be a stepping stone to a better matched job or they may face greater barriers of entry to permanent university graduate level employment which, in turn, would tend to necessitate some level of mismatch for new entrants to gain the necessary experience to obtain a permanent contract. Such a finding is consistent with the predictions of career mobility theory (Sicherman and Galor, 1990). Overeducation is, by contrast, found to be lower the higher the proportion of migrants in the labour market. McGowan et al. (2015)

⁷ The p-value is 0.12

show that countries with a high degree of residential mobility have lower levels of mismatch, as people can travel to different jurisdictions to find matched employment. While McGowan et al. (2015) are referring to residential mobility within countries, migration levels may also proxy residential mobility by indicating that people are willing, and able, to move location to take up matched jobs. This may explain why greater migration is associated with lower overeducation. It is also possible that a higher share of migrant workers is indicative of skill shortages in high skilled posts, or that an inflow of migrant workers could raise the mode of some medium-level occupations, given that, in some countries, such as Ireland, migrants are typically highly educated.⁸ These factors could also explain why a greater share of migrants is associated with lower overeducation.

4.2.1. Young post-secondary (non-tertiary) graduates

The GMM results estimating the determinants of overeducation among young employees aged 15 to 29 with post-secondary qualifications are given in column 2 of Table 3. The lagged dependent variable is positive and significant, demonstrating the persistent nature of the phenomenon. The post-secondary overeducation rate is inversely related to the share of young employees holding post-secondary qualifications, which indicates that youth overeducation at the upper secondary level is not driven upwards by increases in the supply of workers holding post-secondary qualifications. The overeducation rate is also negatively related to the share of young employees holding tertiary level qualifications. As the number of jobs requiring university graduate skills and qualifications increases, this may lower the probability that new entrants holding post-secondary qualifications will be overeducated.

Consistent with previous research, youth post-secondary overeducation is countercyclical and is lower during periods of rising per-capita incomes. However, post-secondary youth overeducation is positively related to the labour force participation rate. Finally, post-secondary overeducation is lower in countries with a higher share employed in high versus low skilled occupations.

The coefficients in the GMM model (Table 3) are not easily interpretable and need to be transformed in order to calculate the long-run marginal effects of the significant variables. The relevant long-run marginal effects are given in Table 4. They show that for every one percent increase in the share of young employees with ISCED 5+ qualifications, the ISCED 5+ youth

⁸ This has certainly been the case in Ireland. For Instance, Barrett et al. (2012) report that 54 per cent of migrants from EU accession states in Ireland held a third level qualification compared to 45 per cent of Irish Nationals.

overeducation rate falls by almost one third of a percent. With respect to ISCED 3-4 overeducation, a one percent increase in the share of young employees holding this medium-level of education will generate a 1.14 percent reduction in long-run overeducation amongst this group. Other points to note are that cyclical effects are small, with a one percent increase in per capita GDP leading to a 0.1 percent reduction in long-run overeducation among young employees with ISCED level 3-4 qualifications.

< Insert Table 4 >

In terms of the model diagnostics, while the model for overeducation at ISCED 5 passes both the A-Bond test for autocorrelation and the Sargan test for overidentification at conventional levels, the ISCED 3 /4 overeducation model fails the Sargan test. On examining the data, we found that while graduate overeducation was generally normally distributed, the distribution for post-secondary overeducation was heavily skewed to the left (see appendix, Figures A1 & A2), suggesting that it would be more appropriate to log the dependent variable in this instance. When we re-estimate the model the post-secondary overeducation model with a logged dependent variable, the results are largely unchanged and the model passes the Sargan test for overidentification at a 5 percent level of confidence (Appendix, Table A1).

4.2.2 The role of vocational education

As discussed earlier, it has been suggested that international variations in overeducation rates are linked to the strength of a countries vocational educational sector. It is plausible to suggest that, for instance, individuals are less likely to progress to higher education in countries with more developed vocational pathways and that this, in turn, will have a dampening impact on graduate overeducation rates. Furthermore, given that many forms of vocational training are sponsored by employers and involve work-based learning, a large vocational sector will also result in lower rates of overeducation among new entrants with ISCED level 3 / 4 qualifications, given that this is the educational level at which vocational qualifications are typically obtained. We can explore this hypothesis by exploiting a question in the EULFS waves, that is available from 2014 onwards, that asked individuals whose highest level of education was at ISCED 3 /

4 if they had engaged in vocational training in the previous 4 weeks⁹. Table A2 in Appendix 2 reports the incidence of vocational qualification attainment for young people aged 15 to 29 by country averaged over the 2014 to 2016 period. The average incidence stood at 11.6 per cent, however, this varied substantively; at the upper end of the spectrum rates stood at 21 per cent in France and between 15 and 20 per cent in Austria, Croatia, Czechia, Italy, Netherlands and Slovenia. At the lower end of the spectrum a zero incidence was reported in Iceland and Demark¹⁰ with rates of between 1 and 5 percent found for Cyprus, Ireland and Luxembourg.

As the vocational education variable is not available for the entire sample, we re-estimate our model for the 2014 to 2016 period. Given that the central question relates to individuals who have received vocational education in the last 4 weeks, we include in our model a measure that relates to the share of 15 to 29 year olds in receipt of vocational education¹¹. Furthermore, as our data relates to a much more restricted sample, we estimate a fixed effects model as opposed to adopting a GMM estimator. On the basis that the vocational education variable will be highly correlated with the share of young workers qualified to ISCED 3 / 4 level we exclude this control from our model. Finally, given that the vocational share of education will tend to be correlated with respect to occupational structure¹², these controls are also omitted. The results from the model are presented in table 5 and they provide no support for the hypothesis that countries with more developed vocational education pathways have lower rates of either graduate or post-secondary youth overeducation. The finding with respect to third level is perhaps not surprising given our central result that many European labour markets seemed to have sufficient capacity to absorb rapid increases in third-level supply without experiencing increases in graduate overeducation. It is also worth noting that the graduate and upper/post-secondary employment shares are not significant in this model. These results should be interpreted with some caution, as the shorter time period of 2014-2016 will capture only a limited amount of change in the educational structure of employment that has occurred over recent decades. Moreover, the sample sizes are relatively small.

⁹ Specifically, the vocational education variable relates only to those who answered that they were a student or an apprentice in regular education during the last 4 weeks and they answered that this level of education or training was ISCED 3 or ISCED 4.

¹⁰ It is unclear why zero rates have been reported here, however, this is most likely to reflect peculiarities of accreditation systems in these countries.

¹¹ As opposed to the share of full-time employees aged 15 to 29 in receipt of vocational education.

¹² For instance many vocational programmes will be sponsored by employers, thus countries with more employment in sectors such as construction and manufacturing are, almost by definition, likely to have more developed educational pathways.

< Insert Table 5 >

4.2.3 *Robustness check*

One potential criticism of the approach adopted in this paper is that the fall in graduate overeducation is simply due to rising educational participation pushing up modal educational levels. In such cases, falling overeducation rates could reflect a statistical artefact as opposed to changes in occupational employment growth being skewed towards occupations requiring higher levels of schooling. In order to examine the possibility that the results are being driven by changes in “technical overeducation”, we examine the extent to which occupational modes have changed over the sample period and re-estimate a model on occupations with unchanged modes. Given that the occupational classification system used in the LFS changed after 2010 we cannot include data from 2011 onwards; the starting point for the data is Q3 2003 given that this is the first time point with complete information for all countries. Our analysis indicates that the modal level of education remained unchanged for 80 per cent of the occupations within countries. The incidence of modal changes at the country level are reported in Table A3. Across all countries in our data, we had 793 occupational groupings and, of these, the occupational mode changed in 154 between 2003 and 2010, equating to 19.4 per cent of the occupations. While less than 20 per cent of occupational modes changed over the period, differences did occur at country level. Over two-thirds of occupational modes changed in Bulgaria and Sweden, with the figure for Ireland standing at 44 percent; conversely, the rate of change was zero in France and less than 5 percent in Germany and Czechia. These results would suggest that the empirical approach to measuring overeducation could provide a potentially distorted picture when applied at country level in some instances.

When we restrict the sample to include only occupations within countries with constant occupational modes between 2003 and 2010, we find that the incidence of youth overeducation falls only marginally, demonstrating that the impacts of “technical overeducation” is likely to be minimal. In the full sample the ISCED 5 and ISCED 3 / 4 average incidence of overeducation during 2003 to 2010 stood at 49.6 and 16.7 percent respectively; when the sample was restricted to occupations with unchanging educational modes, the ISCED 5 and ISCED 3 / 4 averages stood at 45.9 and 14.1 percent respectively.

When running the GMM models, for comparability, we firstly re-estimate our baseline model on the restricted time period and show the results in Table A4. We then estimate the model for the sample restricted to occupations with constant modes (Table A5). While not significant at conventional levels, the ISCED 5 own share is negative for this period. This is consistent with our main finding that increased educational attainment is not associated with increasing overeducation. The key point is that the coefficients in both tables A4 and A5 are similar in relation to their signs and magnitudes. This indicates that our central results are unaffected by declines in overeducation related to increases in occupational modes arising from rapid increases in educational attainment.

5.0 Conclusions

A vast literature exists examining the determinants and consequences of overeducation. However, the focus of existing research has been primarily on overeducation of tertiary graduates within specific countries, or a group of countries, at a static point in time. In this paper, we fill a gap in the literature by constructing a quarterly time series dataset of overeducation rates for 30 European countries from 2000 to 2016. Rather than restricting our analysis to university graduate overeducation, we also contribute to the existing research by studying overeducation rates among workers educated to post-secondary (non-tertiary) level. Moreover, the study focuses on young workers, who have responded the most to EU higher education expansion policies of recent years, including the key EU2020 benchmark targets of ensuring that at least 40% of people aged 30-34 will have completed some form of higher education.¹³

Our analysis shows a rapid increase in tertiary educational attainment among young people, increasing from 11.5 percent in 2000 to 18 percent in 2016. However, despite this, overeducation rates among young university graduate level employees decreased by a third over the same period, from approximately 60 percent in 2000 to around 40 percent in 2016. These descriptive results indicate that EU labour markets have been able to absorb increases in graduate labour supply in recent decades. Our econometric techniques reinforce these findings, showing that tertiary graduate overeducation rates are inversely related to the share of employees educated to tertiary level. This is an important finding, as it contradicts the widely held assumption that overeducation is due to expanding higher educational supply outstripping

¹³ https://ec.europa.eu/education/policies/european-policy-cooperation/et2020-framework_en

the demand for graduate labour. Similarly, the percentage of employees educated to post-secondary (non-tertiary) level fell from 47.2 percent in 2000 to 44.2 percent in 2016. The average overeducation rate among this group also fell from 13.8 per cent to 12.6 per cent over the same period, with our empirical estimates showing that overeducation rates of medium-qualified workers tend to be more responsive to labour supply and demand compositional effects. We find no evidence to support the view that overeducation rates among young people are lower in countries with more developed vocational education pathways.

Overall, the results of our study highlight the importance of a number of alternative policies, other than reliance on educational expansion, for mitigating the incidence of youth overeducation in EU countries. Similar to the conclusions of McGuinness et al. (2018a), gender equality policies and policies to spur a greater supply of high-skilled job openings (e.g. larger R&D investments), which can absorb the greater influx of skills from rising labour force participation rates, may lead to declining overeducation especially among medium-level young graduates. Policies that can further stimulate the reallocation of labour across jobs may also contribute to young graduates finding for themselves suitably matched jobs during the early stages of their careers. Finally, policies aimed at increasing the growth in quality in line with educational attainment and those aimed at removing informational asymmetries as people transition from education to the labour market will also help to reduce the problem.

Figures and Tables

Figure 1

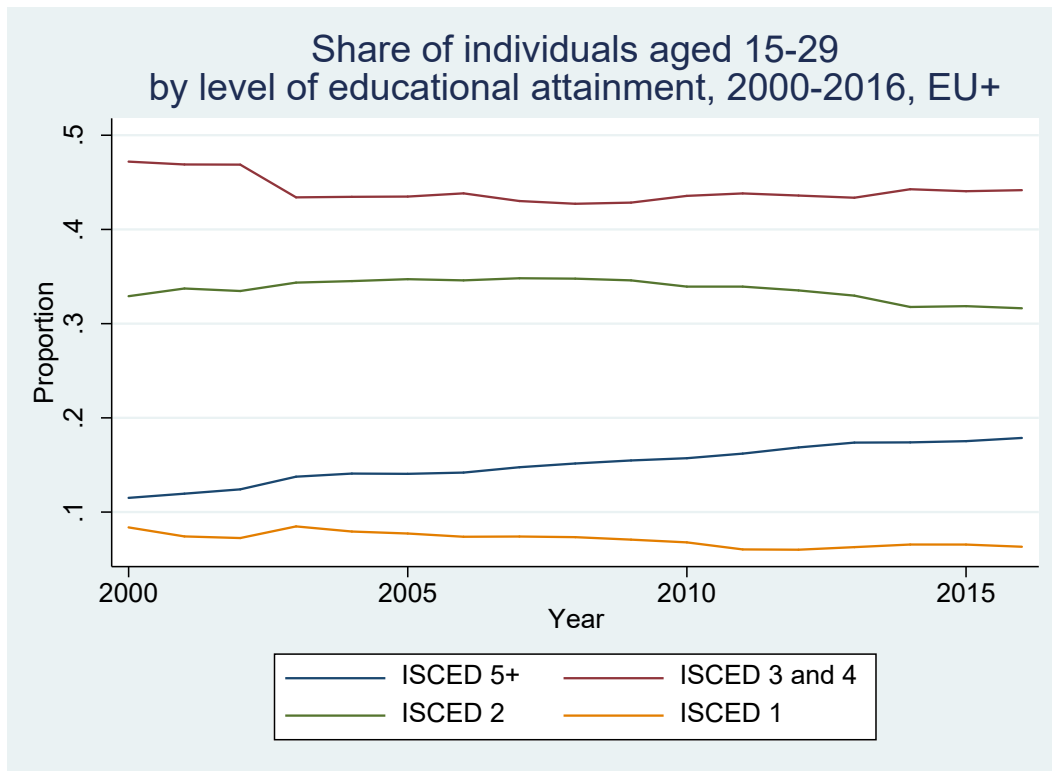


Figure 2

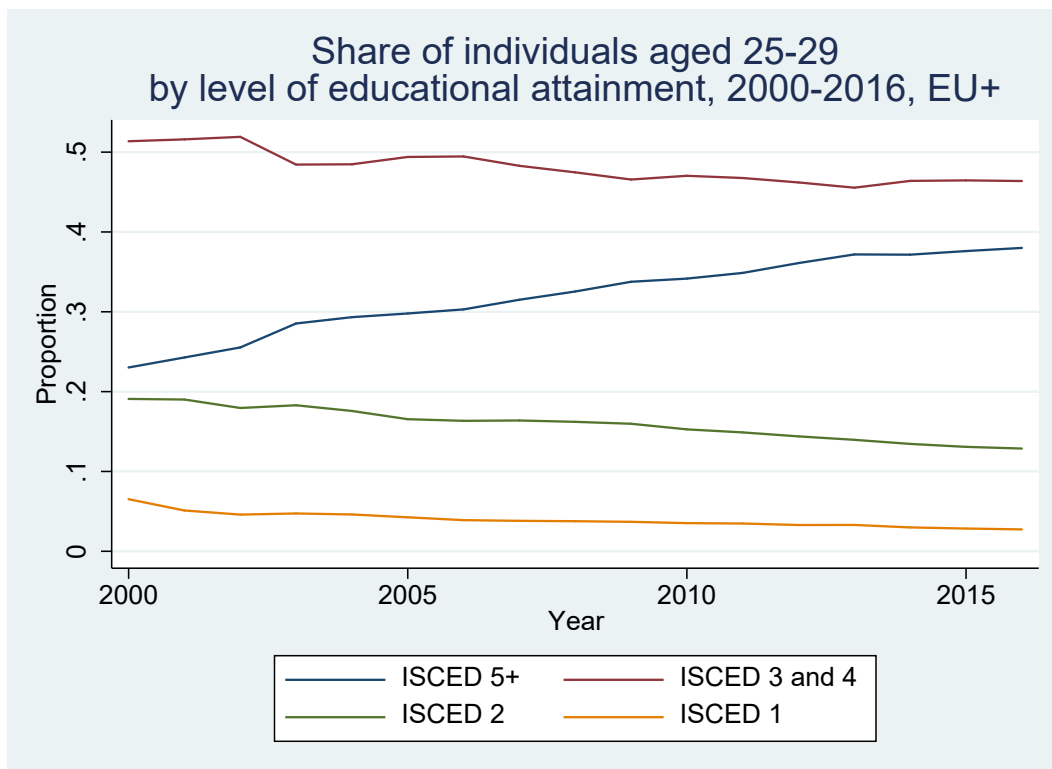


Figure 3

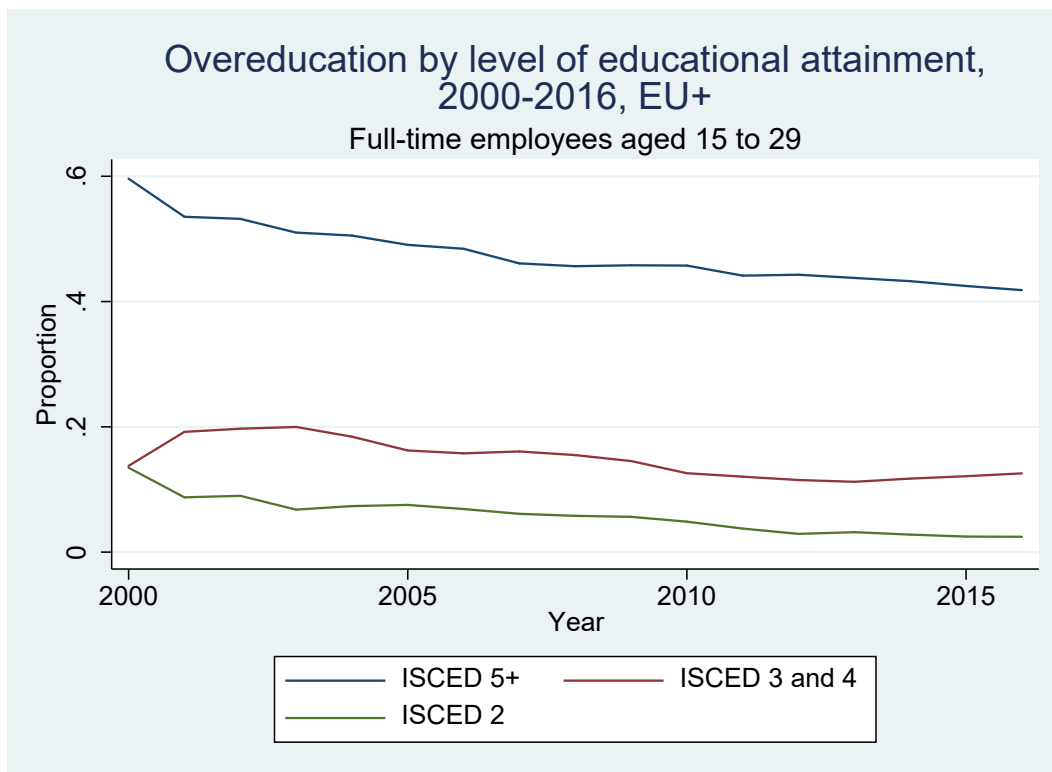


Table 1: Youth educational attainment and overeducation rates for ISCED 3-4 level graduates in 2000 and 2016

| Year | 2000 | | 2016 | |
|-----------------|---------------------------|----------------------------|---------------------------|----------------------------|
| Country | Proportion in ISCED 3 & 4 | Over-Education ISCED 3 & 4 | Proportion in ISCED 3 & 4 | Over-Education ISCED 3 & 4 |
| Austria | 0.544 | 0.014 | 0.464 | 0.049 |
| Belgium | 0.425 | 0.087 | 0.413 | 0.055 |
| Bulgaria | 0.506 | . | 0.476 | 0.024 |
| Croatia | . | . | 0.582 | 0.016 |
| Cyprus | 0.428 | 0.260 | 0.398 | 0.034 |
| Czechia | 0.651 | 0 | 0.516 | 0 |
| Denmark | 0.493 | 0.134 | 0.337 | 0.003 |
| Estonia | 0.487 | 0.002 | 0.428 | 0.102 |
| Finland | 0.469 | 0 | 0.458 | 0.012 |
| France | 0.399 | 0 | 0.425 | 0 |
| Germany | 0.609 | 0 | 0.492 | 0.198 |
| Greece | 0.513 | 0.158 | 0.514 | 0.238 |
| Hungary | 0.573 | 0.050 | 0.493 | 0.374 |
| Iceland | 0.294 | 0.633 | 0.358 | 0.372 |
| Ireland | 0.356 | 0.262 | 0.416 | 0.119 |
| Italy | 0.480 | 0.317 | 0.456 | 0.338 |
| Latvia | 0.509 | 0 | 0.439 | 0.081 |
| Lithuania | 0.449 | 0 | 0.395 | 0.187 |
| Luxembourg | 0.524 | 0.149 | 0.314 | 0.131 |
| Netherlands | 0.433 | 0.101 | 0.386 | 0.009 |
| Norway | 0.544 | 0 | 0.341 | 0.032 |
| Poland | 0.541 | 0.001 | 0.458 | 0.034 |
| Portugal | 0.225 | 0.479 | 0.369 | 0.497 |
| Romania | 0.557 | 0.001 | 0.475 | 0.025 |
| Slovak Republic | 0.670 | 0 | 0.500 | 0.019 |
| Slovenia | 0.507 | 0.116 | 0.518 | 0.010 |
| Spain | 0.316 | 0.699 | 0.339 | 0.651 |
| Sweden | 0.466 | 0 | 0.439 | 0.100 |
| Switzerland | 0.521 | 0 | 0.440 | 0.041 |
| UK | 0.500 | 0.194 | 0.541 | 0.002 |
| | | | | |
| Total Average | 0.472 | 0.138 | 0.442 | 0.126 |

Note: Croatia is not available for the year 2000. Full-time variable is missing for Bulgaria in 2000. The proportion in ISCED refers to individuals ages 15-29 and the overeducation rates refer to full-time employees aged 15 to 29.

Table 2: Youth educational attainment and overeducation rates for ISCED 5+ level graduates in 2000 and 2016

| Time | 2000 | | 2016 | |
|-----------------|------------------------|-------------------------|------------------------|-------------------------|
| Country | Proportion in ISCED 5+ | Over-Education ISCED 5+ | Proportion in ISCED 5+ | Over-Education ISCED 5+ |
| Austria | 0.057 | 0.744 | 0.230 | 0.608 |
| Belgium | 0.184 | 0.486 | 0.230 | 0.198 |
| Bulgaria | 0.073 | . | 0.121 | 0.436 |
| Croatia | . | . | 0.142 | 0.482 |
| Cyprus | 0.160 | 0.598 | 0.305 | 0.397 |
| Czechia | 0.045 | 0.621 | 0.154 | 0.534 |
| Denmark | 0.131 | 0.508 | 0.161 | 0.263 |
| Estonia | 0.094 | 0.601 | 0.144 | 0.342 |
| Finland | 0.132 | 0.399 | 0.127 | 0.290 |
| France | 0.193 | 0.639 | 0.204 | 0.630 |
| Germany | 0.096 | 0.596 | 0.130 | 0.329 |
| Greece | 0.089 | 0.562 | 0.157 | 0.494 |
| Hungary | 0.059 | 0.373 | 0.102 | 0.486 |
| Iceland | 0.096 | 0.385 | 0.149 | 0.456 |
| Ireland | 0.142 | 0.687 | 0.237 | 0.310 |
| Italy | 0.041 | 0.682 | 0.106 | 0.516 |
| Latvia | 0.062 | 0.434 | 0.174 | 0.348 |
| Lithuania | 0.173 | 0.383 | 0.171 | 0.281 |
| Luxemburg | 0.128 | 0.255 | 0.164 | 0.217 |
| Netherlands | 0.116 | 0.511 | 0.182 | 0.254 |
| Norway | 0.227 | 0.565 | 0.219 | 0.440 |
| Poland | 0.049 | 0.289 | 0.191 | 0.346 |
| Portugal | 0.054 | 0.361 | 0.150 | 0.399 |
| Romania | 0.040 | 0.273 | 0.100 | 0.395 |
| Slovak Republic | 0.045 | 0.562 | 0.175 | 0.610 |
| Slovenia | 0.094 | 0.542 | 0.187 | 0.504 |
| Spain | 0.233 | 0.695 | 0.236 | 0.333 |
| Sweden | 0.201 | 0.465 | 0.232 | 0.412 |
| Switzerland | 0.112 | 0.60 | 0.177 | 0.340 |
| UK | 0.197 | 0.647 | 0.267 | 0.506 |
| | | | | |
| Total Average | 0.115 | 0.596 | 0.179 | 0.418 |

Note: Croatia is not available for the year 2000. Full-time variable is missing for Bulgaria in 2000. The proportion in ISCED refers to individuals ages 15-29 and the overeducation rates refer to full-time employees aged 15 to 29.

Table 3: Determinants of Over-Education by education attainment level, GMM estimates, 2000-2016, 29 European countries

| VARIABLES | (1) Over-Education ISCED 5+ | (2) Over-Education ISCED 3 & 4 |
|--|--------------------------------|-----------------------------------|
| Share in ISCED 3 & 4 | 0.12*** (0.045) | -0.61*** (0.108) |
| Share in ISCED 5+ | -0.14*** (0.054) | -0.17*** (0.070) |
| % Female Workers | 0.03 (0.294) | 0.22 (0.231) |
| % of Migrant Workers | -0.23** (0.093) | -0.05 (0.125) |
| % of Temporary Workers | 0.13 (0.130) | -0.05 (0.111) |
| % of Part-time Workers | -0.05 (0.468) | -0.20 (0.158) |
| Unemployment Rate | -0.11 (0.137) | -0.25 (0.177) |
| % Employed in Public Admin | 0.10 (0.431) | 0.58 (0.539) |
| % Employed in Sales | 0.17 (0.402) | -0.14 (0.284) |
| % Employed in Manufacturing | 0.17 (0.240) | -0.14 (0.240) |
| Labour Force Participation | -0.13 (0.159) | 0.42*** (0.132) |
| Research and Development | 0.00 (0.021) | -0.02 (0.014) |
| Log GDP per Capita | -0.01 (0.036) | -0.06** (0.026) |
| Ratio of High (2,3) to Low (7, 8,9) SOC | -0.00 (0.023) | -0.05*** (0.018) |
| Lagged Overeducation 5+ | 0.55*** (0.039) | |
| Lagged Overeducation 3 & 4 | | 0.47*** (0.079) |
| Constant | 0.35 (0.356) | 0.84*** (0.320) |
| Observations | 1,552 | 1,492 |
| Number of countries | 29 | 29 |
| e(F) | 0 | 0 |
| A-Bond Test AR(1), $p(H_0: \text{no autocorrelation})$ | 0.000 | 0.0008 |
| A-Bond Test AR(2), $p(H_0: \text{no autocorrelation})$ | 0.192 | 0.132 |
| Sargan Test, $p(H_0: \text{overidentifying restrictions are valid})$ | 0.156 | 0.000 |

*Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ Note Switzerland is not included in the models due to missing data on the research and development variable.*

Table 4: Estimates of Long-Run Effects derived from GMM Model

| VARIABLES | (1) Over-Education ISCED 5+ | (2) Over-Education ISCED 3 & 4 |
|---------------------------------|--------------------------------|--------------------------------------|
| Share in ISCED 5+ | -0.306*** (0.083) | -0.327*** (0.053) |
| Share in ISCED 3 & 4 | 0.260*** (0.062) | -1.14*** (0.040) |
| % of Migrants | -0.502** (0.198) | |
| Labour Force Participation Rate | | 0.787*** (0.141) |
| Log GDP per capita | | -0.111*** (0.030) |
| Ratio of High to Low SOC | | -0.098*** (0.013) |

*Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ Note Switzerland is not included in the models due to missing data on research and development. We only calculate long run effects for variables which are significant at the 5% level or below. The long-run effect of the covariate is usually defined to be the sum of the current (and lagged coefficients, if used) divided by 1 minus the sum of the lagged coefficients on the dependent variable*

Table 5: Determinants of Over-Education by education attainment level controlling for vocational education share, Fixed Effects estimates, 2014-2016, 27 European countries

| VARIABLES | (1) Over-Education ISCED 5+ | (2) Over-Education ISCED 3 & 4 |
|---|--------------------------------|-----------------------------------|
| Share in ISCED 5+ | -0.22 (0.165) | 0.15 (0.108) |
| Share in Vocational Education | 0.40 (0.398) | 0.04 (0.308) |
| % Female Workers | -0.05 (0.515) | 0.86** (0.403) |
| % of Migrant Workers | -0.09 (0.384) | -0.06 (0.341) |
| % of Temporary Workers | 0.68* (0.377) | 0.25* (0.144) |
| % of Part-time Workers | 0.56 (0.540) | -0.82*** (0.176) |
| Unemployment Rate | -0.14 (0.435) | -0.10 (0.203) |
| Labour Force Participation | 0.02 (0.336) | -0.06 (0.170) |
| Research and Development | 0.04 (0.045) | 0.02 (0.020) |
| Log GDP per Capita | 0.01 (0.034) | -0.02 (0.020) |
| Ratio of High (2,3) to Low (7, 8,9) SOC | -0.04 (0.029) | -0.01 (0.015) |
| Constant | 0.25 (0.386) | 0.07 (0.241) |
| Observations | 243 | 243 |
| R-squared | 0.12 | 0.20 |
| Number of countries | 27 | 27 |
| e(F) | 0.00905 | 3.40e-09 |

*Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. We are missing Czechia which has 0 rates of overeducation at ISCED level 3 & 4 over the period 2014-2016. We also are missing Norway and Iceland which has zero rates of vocational education attainment over 2014-2016.*

References

- Barrett, A., McGuinness, S. and O'Brien, M. (2012). The immigrant earnings disadvantage across the earnings and skills distributions: The case of immigrants from the EU's new member states. *British Journal of Industrial Relations*, 50 (3), pp. 457-481.
- Baert, S., Cockx, B. and Verhaest, D. (2013) Overeducation at the start of the career: Stepping stone or trap? *Labour Economics*, 25, pp.123-140.
- Budría, S. and Moro-Egido, A.I. (2018) Qualification and skill mismatches: Europe in a cross-national perspective. *Cuadernos Económicos de ICE*, 95, pp.151-188.
- Carroll, D. and Tani, M. (2015). Job search as a determinant of graduate over-education: Evidence from Australia. *Education Economics*, 23 (5), pp.631-644.
- Chevalier, A. and Lindley, J. (2009) Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172 (2), pp.307-337.
- Croce, G. and Ghignoni, E. (2012) Demand and supply of skilled labour and overeducation in Europe: A country-level analysis. *Comparative Economic Studies*, 54 (2), pp.413-439.
- Davia, M.A., McGuinness, S. and O'Connell, P.J. (2017) Determinants of regional differences in rates of overeducation in Europe. *Social Science Research*, 63, pp.67-80.
- Lessaer, B., Pasimeni, P., Pouliakas, K. et al. (2015), 'Supporting skills development and matching in the EU', Chapter III.1 in European Commission, *Employment and Social Developments in Europe 2015*, Luxembourg: Publications Office of the European Union.
- Li, I.W. and Miller, P.W. (2015) Overeducation and earnings in the Australian graduate labour market: An application of the Vahey model. *Education Economics*, 23 (1), pp.63-83.
- Mavromaras, K. and McGuinness, S. (2012) Overskilling dynamics and education pathways. *Economics of Education Review*, 31 (5), pp.619-628.
- McGowan, M.A., Adalet, M. and Andrews, D. (2015) Skill mismatch and public policy in OECD countries. OECD Working Paper 28, Organisation for Economic Co-operation and Development, Paris.
- McGuinness, S. (2006) Overeducation in the labour market. *Journal of Economic Surveys*, 20 (3), pp.387-418.

- McGuinness, S., Bergin, A. and Whelan, A. (2018) Overeducation in Europe: Trends, convergence, and drivers. *Oxford Economic Papers*, 70 (4), pp.994-1015.
- McGuinness, S., Pouliakas, K. and Redmond, P. (2018) Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys*, 32 (4), pp.985-1015.
- Ordine, P. and Rose, G. (2017) Too many graduates? A matching theory of educational mismatch. *Journal of Human Capital*, 11 (4), pp. 423-446.
- Pouliakas (2012), 'The skill mismatch challenge in Europe', Chapter 6 in European Commission, *Employment and Social Developments in Europe 2012*, Luxembourg: Publications Office of the European Union.
- Quintini, G. (2011) Over-qualified or under-skilled: A review of existing literature. OECD Social, Employment and Migration Working Paper 121, Organisation for Economic Co-operation and Development, Paris.
- Sicherman, N. and Galor, O. (1990) A theory of career mobility. *Journal of Political Economy*, 98 (1), pp.169-192.
- Verhaest, D. and Van der Velden, R. (2013) Cross-country differences in graduate overeducation. *European Sociological Review*, 29 (3), pp.642-653.

Appendix

Figure A1

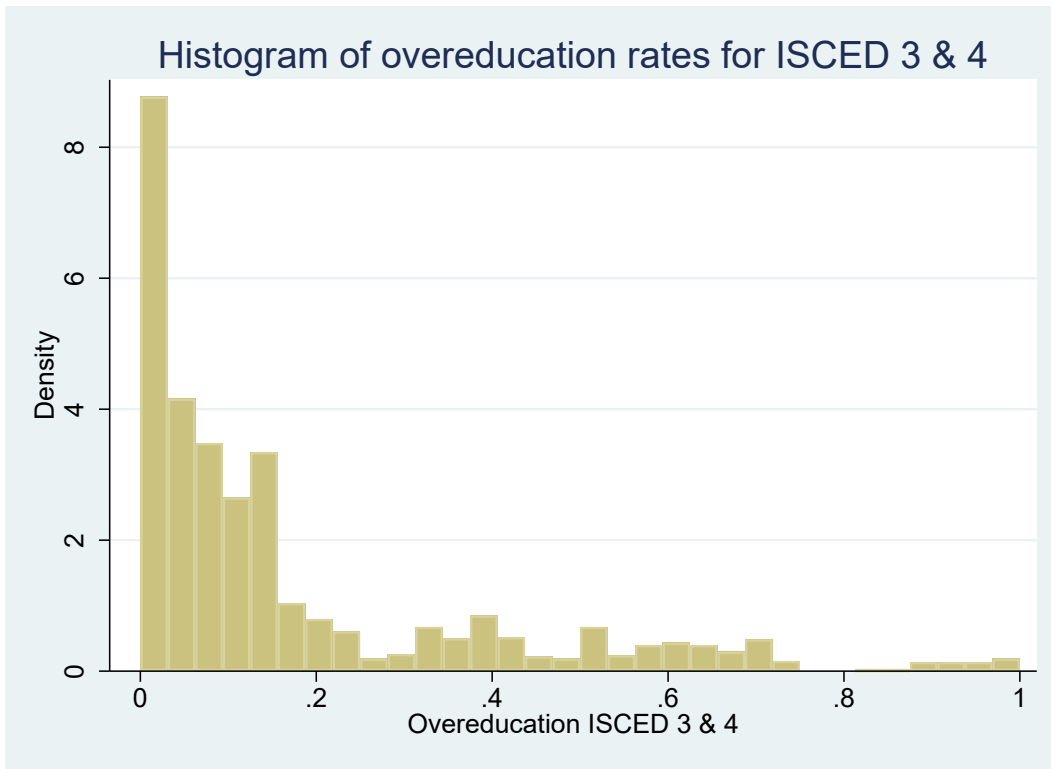


Figure A2

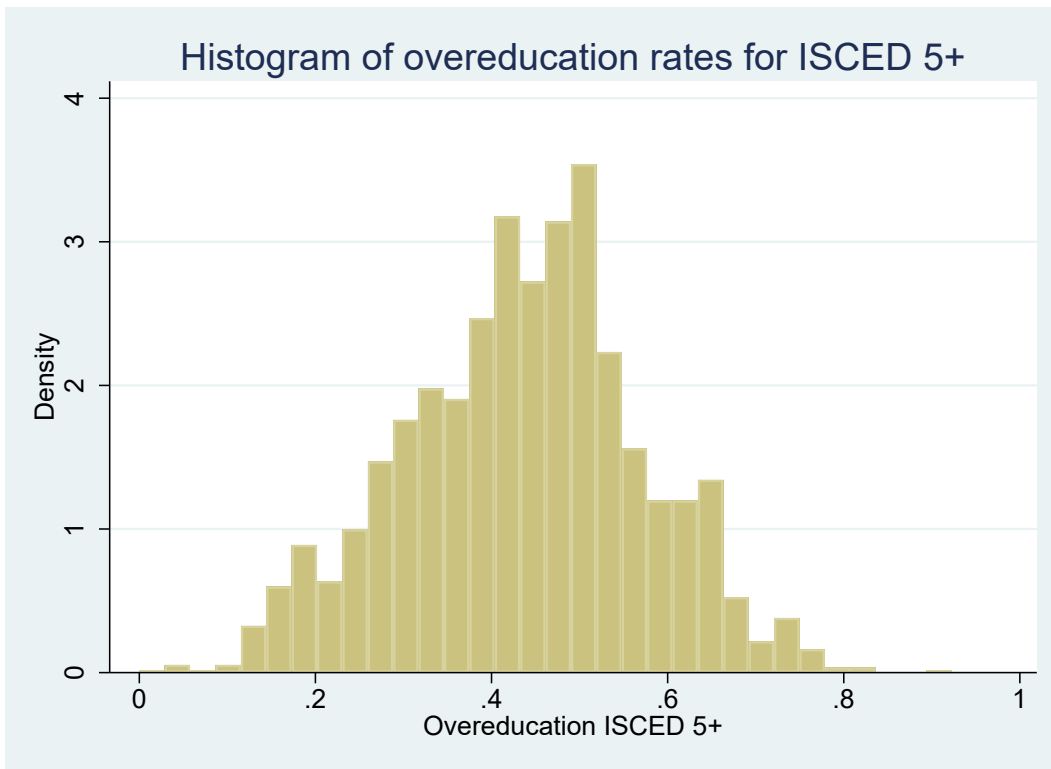


Table A1: Determinants of Over-Education by education attainment level, GMM estimates, 2000-2016, 29 European countries (logged dependent variable)

| VARIABLES | (2) Log Over-Education ISCED 3 & 4 |
|--|---------------------------------------|
| Share in ISCED 3 & 4 | -2.43** (0.948) |
| Share in ISCED 5+ | -0.50 (0.775) |
| % Female Workers | 4.93** (2.355) |
| % of Migrant Workers | -1.10 (1.027) |
| % of Temporary Workers | -1.34 (0.889) |
| % of Part-time Workers | 0.22 (1.041) |
| Unemployment Rate | -2.34*** (0.844) |
| % Employed in Public Admin | 4.75 (3.663) |
| % Employed in Sales | -4.27 (3.668) |
| % Employed in Manufacturing | -1.78 (2.625) |
| Labour Force Participation | 2.97** (1.159) |
| Research and Development | -0.20* (0.105) |
| Log GDP per Capita | -0.53*** (0.163) |
| Ratio of High (2,3) to Low (7, 8,9) SOC | -0.58*** (0.190) |
| Lagged Overeducation 5+ | |
| Lagged Overeducation 3 & 4 | 0.56*** (0.075) |
| Constant | 3.70 (2.325) |
| Observations | 1,372 |
| Number of countries | 29 |
| e(F) | 0 |
| A-Bond Test AR(1), $p(H_0: \text{no autocorrelation})$ | 0.002 |
| A-Bond Test AR(2), $p(H_0: \text{no autocorrelation})$ | 0.874 |
| Sargan Test, $p(H_0: \text{overidentifying restrictions are valid})$ | 0.064 |

*Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ Note Switzerland is not included in the models due to missing data on the research and development variable.*

Table A2: Youth vocational educational attainment aged 15 to 29 for 30 EU countries, 2014-2016

| Country | Proportion with Vocational Qualifications as Highest ISCED Level |
|-----------------|---|
| Austria | 0.186 |
| Belgium | 0.112 |
| Bulgaria | 0.066 |
| Croatia | 0.183 |
| Cyprus | 0.032 |
| Czechia | 0.199 |
| Denmark | 0.065 |
| Estonia | 0.084 |
| Finland | 0.117 |
| France | 0.214 |
| Germany | 0.127 |
| Greece | 0.050 |
| Hungary | 0.052 |
| Iceland | 0 |
| Ireland | 0.030 |
| Italy | 0.164 |
| Latvia | 0.088 |
| Lithuania | 0.053 |
| Luxemburg | 0.016 |
| Netherlands | 0.163 |
| Norway | 0 |
| Poland | 0.142 |
| Portugal | 0.099 |
| Romania | 0.125 |
| Slovak Republic | 0.115 |
| Slovenia | 0.157 |
| Spain | 0.069 |
| Sweden | 0.071 |
| Switzerland | 0.053 |
| UK | 0.065 |
| | |
| Total Average | 0.115 |

Table A3: Occupations that have ISCED Mode Changes over 2003 to 2010, by Country

| Country | Number of Occupations with Changing ISCED Mode | Total Number of occupations | Percent of Total Occupations |
|-----------------|--|-----------------------------|------------------------------|
| Austria | 6 | 27 | 0.222 |
| Belgium | 3 | 26 | 0.115 |
| Bulgaria | 18 | 27 | 0.667 |
| Croatia | 4 | 27 | 0.148 |
| Cyprus | 9 | 27 | 0.333 |
| Czechia | 1 | 27 | 0.037 |
| Denmark | 6 | 31 | 0.193 |
| Estonia | 2 | 27 | 0.074 |
| Finland | 3 | 27 | 0.111 |
| France | 0 | 29 | 0 |
| Germany | 1 | 27 | 0.037 |
| Greece | 7 | 27 | 0.259 |
| Hungary | 6 | 27 | 0.222 |
| Iceland | 4 | 26 | 0.154 |
| Ireland | 12 | 27 | 0.444 |
| Italy | 4 | 26 | 0.154 |
| Latvia | 5 | 27 | 0.185 |
| Lithuania | 3 | 27 | 0.111 |
| Luxemburg | 7 | 27 | 0.259 |
| Netherlands | 2 | 29 | 0.069 |
| Norway | 3 | 26 | 0.115 |
| Poland | 2 | 27 | 0.074 |
| Portugal | 2 | 27 | 0.074 |
| Slovak Republic | 2 | 27 | 0.074 |
| Slovenia | 4 | 27 | 0.148 |
| Spain | 6 | 27 | 0.222 |
| Sweden | 19 | 27 | 0.704 |
| Switzerland | 6 | 33 | 0.182 |
| UK | 7 | 27 | 0.259 |
| | | | |
| Total Average | 5.3 | 29 | 0.194 |

Note: Romania has all occupations coded as either 10, 20, 30, etc in 2003 and as 11, 12,..21,22, etc. in 2013 so there were no occupations with the same code over the period.

Table A4: Determinants of Over-Education by education attainment level, GMM estimates varying modes, 2000-2010, 29 European countries

| VARIABLES | (1) Over-Education ISCED 5+ | (2) Over-Education ISCED 3 & 4 |
|--|-----------------------------------|--------------------------------------|
| Share in ISCED 3 & 4 | 0.21*** (0.057) | -0.78*** (0.112) |
| Share in ISCED 5+ | -0.11 (0.070) | -0.26*** (0.074) |
| % Female Workers | 0.13 (0.495) | -0.00 (0.262) |
| % of Migrant Workers | -0.07 (0.101) | 0.08 (0.158) |
| % of Temporary Workers | 0.05 (0.223) | -0.01 (0.153) |
| % of Part-time Workers | 0.34 (0.608) | -0.51*** (0.164) |
| Unemployment Rate | -0.24 (0.176) | -0.01 (0.165) |
| % Employed in Public Admin | -0.15 (0.679) | -0.81* (0.449) |
| % Employed in Sales | 0.23 (0.552) | -0.49 (0.401) |
| % Employed in Manufacturing | 0.39 (0.332) | -0.14 (0.316) |
| Labour Force Participation | 0.23 (0.317) | -0.21 (0.187) |
| Research and Development | -0.02 (0.033) | 0.01 (0.026) |
| Log GDP per Capita | -0.06 (0.049) | -0.04 (0.034) |
| Ratio of High (2,3) to Low (7, 8,9) SOC | -0.01 (0.048) | -0.06* (0.031) |
| Lagged Overeducation 5+ | 0.34*** (0.060) | |
| Lagged Overeducation 3 & 4 | | 0.32*** (0.066) |
| Constant | 0.53 (0.534) | 1.43*** (0.391) |
| Number of countries | 29 | 28 |
| Observations | 952 | 906 |

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Note there is very little variation in share with ISCED 3 & 4 in Slovenia over the period 2000-2010 and so that falls out of the model.

Table A5: Determinants of Over-Education by education attainment level, GMM estimates (constant modes) 2000-2010, 29 European countries,

| VARIABLES | (1) Over-Education ISCED 5+ | (2) Over-Education ISCED 3 & 4 |
|--|-----------------------------------|--------------------------------------|
| Share in ISCED 3 & 4 | 0.22*** (0.063) | -0.39** (0.161) |
| Share in ISCED 5+ | -0.06 (0.053) | -0.13*** (0.036) |
| % Female Workers | -0.12 (0.367) | 0.59** (0.277) |
| % of Migrant Workers | 0.10 (0.082) | 0.07 (0.107) |
| % of Temporary Workers | 0.29 (0.211) | 0.16 (0.180) |
| % of Part-time Workers | 0.34 (0.389) | 0.02 (0.124) |
| Unemployment Rate | -0.15 (0.124) | -0.12 (0.112) |
| % Employed in Public Admin | -0.03 (0.492) | 0.14 (0.227) |
| % Employed in Sales | 0.21 (0.394) | -0.54** (0.217) |
| % Employed in Manufacturing | 0.36 (0.338) | -0.24 (0.177) |
| Labour Force Participation | 0.31 (0.345) | -0.09 (0.160) |
| Research and Development | -0.00 (0.024) | -0.00 (0.020) |
| Log GDP per Capita | -0.02 (0.041) | -0.02 (0.024) |
| Ratio of High (2,3) to Low (7, 8,9) SOC | 0.00 (0.033) | -0.08*** (0.025) |
| Lagged Overeducation 5+ | 0.30*** (0.060) | |
| Lagged Overeducation 3 & 4 | | 0.28** (0.114) |
| Constant | 0.11 (0.452) | 0.59*** (0.212) |
| Number of countries | 29 | 28 |
| Observations | 952 | 900 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A6: Data Description

| Model Covariate Name | Description |
|--|---|
| Share in ISCED 3 & 4 | Proportion of full-time employees aged 15 to 29 with highest educational attainment of ISCED 3 or 4 |
| Share in ISCED 5+ | Proportion of full-time employees aged 15 to 29 with highest educational attainment of ISCED 5 or higher |
| % Female Workers | Proportion of full-time employees that are female |
| % of Temporary Workers | Proportion of full-time employees that are temporary workers |
| % of Part-time Workers | Proportion of employees that are part-time workers |
| % of Migrant Workers | Proportion of full-time employees that are migrant workers |
| Unemployment Rate | Overall unemployment rate |
| % Employed in Public Admin | Proportion of full-time employees that are employed in public administration sector |
| % Employed in Sales | Proportion of full-time employees that are employed in sales |
| % Employed in Manufacturing | Proportion of full-time employees that are employed in the manufacturing sector |
| Labour Force Participation | Overall labour force participation rate |
| Research and Development | Gross domestic expenditures on Research & Development. This was sourced from the World Bank. Linear interpolation was used to interpolate values for quarters by interpolating between consecutive year values. |
| Log GDP per Capita | Gross domestic product per capita was sourced from Eurostat |
| Ratio of High (2,3) to Low (7,8,9) SOC | Ratio of the proportion of full-time employees employed in SOC 2 and 3 to the proportion of full-time employees employed in SOC 7, 8 and 9 |
| Share in Vocational Education | Proportion of individuals aged 15 to 29 who were in vocational education or training in the last 4 weeks. |