IZA DP No. 10250

Determinants of Regional Differences in Rates of Overeducation in Europe

Maria A. Davia
Seamus McGuinness
Philip J. O'Connell

September 2016

Determinants of Regional Differences in Rates of Overeducation in Europe

Maria A. Davia
Universidad de Castilla – La Mancha

Seamus McGuinness
ESRI, Trinity College Dublin, and IZA

Philip J. O'Connell
UCD Geary Institute

ABSTRACT

Determinants of Regional Differences in Rates of Overeducation in Europe

This paper examines the factors determining variations in spatial rates of overeducation. A quantile regression model has been implemented on a sample of region-yearly data drawn from the EU Survey on Income and Living Conditions (EU-SILC) and several institutional and macroeconomic features captured from other data-sets. Potential determinants of overeducation rates include factors such as labour market risk, financial aid to university students, excess labour demand and institutional factors. We find significant effects both for labour market structural imbalances and institutional factors. The research supports the findings of micro based studies which have found that overeducation is consistent with an assignment interpretation of the labour market.

JEL Classification: C29, I21, J24

Keywords: overeducation, regional variation, mismatch

Corresponding author:

Maria A. Davia
Economics and Business Administration School
Universidad de Castilla – La Mancha
Plaza de la Universidad, s/n
02061, Albacete
Spain
E-mail: mangeles.davia@uclm.es

Determinants of regional Differences in Rates of Overeducation in Europe

1. Introduction

There has been a substantial increase in the educational attainment of populations throughout the advanced industrial societies in recent decades. This has coincided with dramatic growth in the demand for highly educated workers. However, there is concern that the demand for highly educated labour has not kept pace with supply, giving rise to the problem of over-education. Workers are considered overeducated if their qualifications exceed those required for the job (Groot and van den Brink, 2000; McGuinness, 2006). This paper adds to the existing literature by providing an assessment of the potential drivers of overeducation across regions and countries. Unlike limited existing studies that use individual level data to explain cross-country variations in overeducation, we adopt a more aggregate approach that allows us to exploit international and within country regional variations to achieve a more refined assessment of spatial variations in overeducation rates. The analysis also uses an alternative to the standard wage equation framework for assessing theoretical explanations of overeducation and the role of labour market institutions.

Over-education can be costly for individuals, organisations and economies. At the individual level, overeducated workers have been found to earn less than similarly educated workers whose jobs match their qualifications, presumably because a proportion of their investment in education is underutilized and unproductive

(McGuinness and Sloane, 2011; Mavromaras et al , 2009; Bárcena-Martín et al., 2012)[1].

Overeducated workers may also experience lower levels of job-satisfaction (Tsang et al., 1991; Battu et al., 1999). Moreover, less-qualified workers may be displaced and 'bumped down' in the labour market, or into unemployment, by over-educated workers moving into their occupations, particularly in slack labour markets (Battu and Sloane, 2002). At the level of the organisations, there is some evidence to suggest that overeducation may be associated with lower productivity (Tsang, 1987) and higher labour turnover, leading in turn to lost investments in recruitment and training (Tsang et al., 1991; Alba-Ramirez, 1993). At the macroeconomic level, overeducation can entail wastage of investment in education and national output is potentially lower than it could be if the skills of overeducated works were fully utilized.

Literature Review

There is substantial variation in the incidence of overeducation between countries (Di Pietro, 2002; Bárcena-Martín et al., 2012; Croce and Ghignoni, 2012; Verhaest and Van der Velden, 2013). While there has been a surge in the literature on overeducation (see McGuinness (2006) and Sloane (2003) for reviews), the majority of existing work tends to be specific to individual countries and to focus either on measuring the wage effects or on the determinants of country-level education-job mismatch. To date, research to identify the determinants of international differences in rates of overeducation has been limited. Thus, while we know much about the magnitude of overeducation effects on variables such as earnings, job satisfaction and career mobility (Battu, Belfield and Sloane (1999) and Dolton and Vignoles (2000), Peiró et al (2010), McGuinness (2003), McGuinness and Sloane (2011)), there is much less understanding of the structural

---

[1] However, a recent study by Kedir et al (2012) argues that there are no productivity impacts associated with overeducation.

factors that drive the overeducation phenomenon itself. With respect to the very limited work that does exist, Groot and van den Brink (2000), in a meta-analysis, found evidence of a relationship between overeducation and the rate of labour force growth. Hartog (2000, p. 134) suggests that "the strong expansion of participation in education has outpaced the increase in the demanded levels of education". Di Pietro (2002) in a pooled cross-national analysis of aggregate data in 11 countries found that, on the supply side, increases in the educational attainment of the population were associated with higher overeducation, while, on the demand side, increased investment in research and development was associated with lower overeducation. Humburg et al. (2015), in their analysis of graduate overeducation in 17 European countries, found that both field of study and the relationship between supply and demand is important: field-specific education protects against overeducation, and this protective effect is greater in occupations characterized by an excess supply of graduates. Verheast and van der Velden (2013), estimating a multi-level model for a sample of European graduates, found evidence of a role for structural imbalances in both the quantity of skilled workers and their composition in terms of field of study. Croce and Ghignoni (2012), in their pooled model for 26 European countries, found that the ratio of wages of graduates to those of less-qualified workers is associated with graduate overeducation and also that recession leads to overeducation, with graduates accepting jobs requiring less education than they possess. Ghignoni and Verashchagina (2014) explore determinants of individual overeducation risk in 10 European countries taking into account both supply – side and demand – side factors. This paper provides further evidence on the issue, with an assessment of the determinants of international variations in overeducation rates using European data and, in addition to structural factors, assesses the potential

contribution of labour market uncertainty, labour market institutions, education funding mechanisms and migration as determining factors.

The choice of potential covariates to be included in any model explaining cross-country variations in overeducation is not straightforward as there are a number of competing hypotheses on the exact causes of overeducation. Proponents of matching theories of job search (Jovanavic, 1970) suggest that overeducation is largely a consequence of poor information and, over time, workers will realize their error and achieve improved matched through repeated job search. Similarly, theories of career mobility (Rosen, 1972; Sicherman and Galor, 1990) suggest that some workers will deliberately choose mismatch in order to acquire the necessary skills, through on-the-job training and learning that will enable them to achieve more rapid career progression in the future. Therefore, both matching theory and models of career mobility suggest that overeducation is a temporary phenomenon driven by either incomplete information or strategic behaviour and, as such, the phenomenon should be largely unrelated to observable structural factors within an economy. Thurow's Job Competition Model (Thurow, 1975) emphasizes the importance of the characteristics of jobs and argues that workers are allocated to a fixed distribution of jobs with individuals investing in education in order to preserve their place in the jobs queue. Once an individual reaches the top of the queue she is allocated a job and her wage will be predetermined solely by the characteristics of the job in question. Thus, under the Thurow model, overeducation will arise when the number of graduate workers exceeds the number of graduate jobs, thus emphasizing the importance of including variables that reflect any excess supply of educated labour. Assignment models (Sattinger, 1993) also stress the importance of job distribution; however, the job allocation process is no longer a lottery as utility

maximization guides workers to choose certain jobs over others. Thus a central prediction arising from assignment theory is that changes in the distribution of earnings and, by default, overeducation, will be related to both the distribution of jobs and the characteristics of the workforce. Human Capital Theory (HCT) (Becker, 1964; Mincer, 1974) predicts that workers will always earn their marginal product, implying that there will be no under-utilization of human capital in the labour market and that overeducation will not exist in equilibrium. However, McGuinness (2006) points out that overeducation is still consistent with HCT, as a short-run phenomenon, if the stock of educated labour supply rises, as a consequence of higher labour market returns, until such times as firms fully adjust their production processes to accommodate the altered nature of labour supply. Thus, HCT would suggest that any model of overeducation should also include some controls or proxies that capture changes in the rate of return to schooling.

Finally, the role of labour market institutions must be considered as potential drivers of overeducation. To date, empirical papers have focused on testing the consistency of human capital, job competition and assignment theory based on the behaviour of rates of return to required and surplus education. The impact of institutional variables has attracted much less attention, presumably as a consequence of a concentration of research on national datasets within which institutional impacts are treated as fixed effects. Nevertheless, there are some compelling arguments for the impact of institutional factors on education-job matching. Quintini (2011) discusses a number of potential scenarios under which institutional factors can influence the level of mismatch in a region or country. On the one hand, a highly regulated labour market may make it difficult for firms to fire mismatched workers;

alternatively, the prevalence of temporary, fixed term contracts and part-time workers within a more flexible framework may enable firms to circumvent firing regulations to respond to the existence of skill mismatches. Quintini (2011) also argues that rigid wage setting institutions may prevent wage levels from adjusting in response to skill mismatches. In competitive markets wages should adjust in a way that discourages workers from training in areas where there are surplus skills (and lower returns) thus, helping to create a better balance between labour demand and labour supply. Furthermore, McGuinness and Sloane (2011) argue that, for many, overeducation is the product of a conscious trade-off job match between higher wages with other aspects of employment such as job security and an enhanced work-life balance.

The strength of employment protection legislation (EPL) in a country may be associated with lower overeducation because employers may be more risk averse in recruitment leading to better job-skill matching (Gangl, 2004). Strong EPL may also limit the extent to which females are forced to occupationally downgrade in the presence of children, although Verhaest and Van der Velden, (2013) report that EPL impacts were unimportant in explaining inter-country variations in overeducation among a graduate cohort. Adalet McGowan and Andrews (2015), in their analysis of the related concept of skill mismatch[2] in the 22 OECD countries found evidence of the importance of institutional and structural factors that allow for flexibility in labour markets and in reduced barriers to business entry and closure. After controlling for individual and job characteristics, they found that skill mismatch is higher in countries with stronger employment protection

---

[2] Adalet McGowan and Andrews use a measure of skill match that combines workers' self-reported skill match as well as proficiency (literacy) scores collected in the OECD Survey of Adult Skills (PIAAC).

legislation and product market regulation and where bankruptcy laws penalise business closures. They also found that skill mismatch is lower in countries where housing policies do not impede residential (and, thus, geographical) mobility, and in countries with higher rates of participation in life-long learning as well as in those characterised by higher levels of management quality.

To date, the various theoretical frameworks have been tested by assessing the magnitude of the coefficients on required and surplus schooling using a standard wage equation estimated on micro-data, with the evidence generally supporting an assignment interpretation of the labour market and rejecting both the Job Competition Model and HCT (McGuinness, 2006). To an extent, the analyses of cross-country variations in rates of overeducation also allow an alternative framework within which to examine the validity of the various theoretical constructs that are often discussed within the overeducation literature and assess the importance of institutional factors.

2. Materials and Methods

2.1. Data

The data for this study come from the six waves of the European Union Survey on Income and Living Conditions (EU-SILC) collected between 2004 and 2009. A clear advantage of the EU-SILC dataset is its regional geographical component (NUTS1) which provides us with multiple observations for some countries, thus generating a workable sample. For each year we have data on a maximum of 28 countries,[3] of

---

[3] Austria, Belgium, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Greece, Hungary, Ireland, Iceland, Italy, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Portugal, Poland, Romania, Sweden, Slovakia, Slovenia and UK.

which regional information is available for eleven,[4] giving us a total of 332 observations (295 of which have information for all explanatory variables in our multivariate analysis) observations over six years[5].

Individuals are defined as being overeducated if their level of attained schooling is at least one level above the mode of their occupation[6]. Therefore, overeducation is measured at the individual level by comparing each respondent's level of education with the modal educational level for their respective 2 digit occupation in their country of residence[7]. Given that our estimated models relate to the region-year, the dependant variables in the models relate to the mean level of overeducation for each gender within each region at each given point in time. This approach to the measurement of overeducation is adopted due to the absence of alternative subjective based measures within EU-SILC, and it is similar to that

---

[4] Austria, Belgium, Bulgaria, Spain, France, Greece, Hungary, Italy, Poland and Sweden. Micro-data in Germany allow to distinguish NUTS 1 only in waves 2004, 2005 and 2006.We have therefore decided to collapse regional information in Germany for the whole observation period

[5] The number of available observations per region and country are displayed in table A1 in the Data Appendix. We only compute in table A1 observations for which there is information in all the explanatory variables of the multivariate analysis. Because of problems with several explanatory variables (namely, financial aid to tertiary education students, employment protection legislation and enrolment rates for 20-24 year olds), we finally do without observations from Cyprus, Malta, Luxembourg and the last three waves of Greece.

[6] This is measured at the two-digit ISCO level. Occupations have been recoded into 27 categories corresponding to the 2-digit classification of ISCO-88, which entails more detailed differences than the one-digit classification in ISCO-88. A more aggregated approach (ie, only 9 categories in 1-digit ISCO-88) would not be satisfactory as it would assume common entry qualifications across highly heterogeneous occupations. For example, if we were to compute overeducation rates for occupations aggregated at 1 digit level, we would have to assume common entry requirement between managers of small enterprises and corporate managers and legislators, as well as between office clerks and persons working in personal and protective services. Table A.3 in the Appendix shows the average occupation-specific overeducation rates. We see that the distribution of overeducation rates is broadly comparable across genders and, as expected, overeducation is largely a consequence of highly educated workers located in lower skilled occupations,

[7] We do not compute individual overeducation at region level because it is unlikely that job-entry conditions differ substantially across regions within countries, and a regional approach would unnecessarily reduce cell sizes.

adopted in previous studies (e.g. Croce et al., 2012; Ghignoni et al., 2014). Moreover, while it has been noted that this realized matches approach tends to show lower rates of overeducation than alternative measures based either on self-assessment of the job-skills match or detailed analysis of job content (Groot et al., (2000), there is no evidence to suggest that a conservative estimate of the incidence of overeducation suffers from cross-national bias. Moreover, previous studies have confirmed that the choice of overeducation measure tends to be of little consequence in terms of the estimated impacts (McGuinness (2006)).

Table 1. Average Overeducation Rates by gender, region and country 2004 - 2009

COUNTRY	REGION	males	females	COUNTRY	REGION	males	females
Austria	AT1	0.228	0.286	Greece	GR1	0.310	0.282
	AT2	0.199	0.267		GR2	0.324	0.206
	AT3	0.218	0.233		GR3	0.320	0.280
average	AT	0.215	0.262		GR4	0.252	0.226
Belgium	BE1	0.120	0.162	*average*	*GR*	*0.301*	*0.249*
	BE2	0.123	0.147	Hungary	HU1	0.135	0.226
	BE3	0.111	0.129		HU2	0.112	0.179
average	*BE*	*0.118*	*0.146*		HU3	0.114	0.194
Bulgaria	BG3	0.093	0.097	*average*	*HU*	*0.120*	*0.200*
	BG4	0.084	0.104	Ireland	*IE0*	*0.352*	*0.288*
average	BG	0.088	0.101	Iceland	*IS*	*0.205*	*0.214*
Czech Republic	CZ0	0.079	0.121	Italy	ITC	0.272	0.280
Germany	*DE*	*0.156*	*0.195*		ITD	0.293	0.312
Denmark	*DK*	*0.148*	*0.122*		ITE	0.289	0.308
Estonia	*EE*	*0.171*	*0.221*		ITF	0.209	0.301
Spain	ES1	0.334	0.272		ITG	0.192	0.293
	ES2	0.394	0.259	*average*	*IT*	*0.251*	*0.299*
	ES3	0.276	0.203	Lithuania	*LT0*	*0.249*	*0.212*
	ES4	0.341	0.251	Latvia	*LV0*	*0.156*	*0.207*
	ES5	0.300	0.219	The Netherlands	*NL*	*0.147*	*0.182*
	ES6	0.300	0.251	Norway	*NO*	*0.073*	*0.087*
	ES7	0.310	0.236	Portugal	*PT*	*0.285*	*0.265*
Average	*ES*	*0.322*	*0.242*	Poland	PL1	0.101	0.193
Finland	*FI*	*0.062*	*0.070*		PL2	0.087	0.172
France	FR1	0.130	0.092		PL3	0.091	0.176
	FR2	0.091	0.086		PL4	0.088	0.172
	FR3	0.095	0.092		PL5	0.080	0.202
	FR4	0.128	0.094		PL6	0.080	0.166
	FR5	0.107	0.095	*Average*	*PL*	*0.088*	*0.180*
	FR6	0.110	0.105	Sweden	SE0	0.142	0.120
	FR7	0.088	0.088		SE1	0.161	0.149
	FR8	0.126	0.126		SE2	0.141	0.122
average	*FR*	*0.109*	*0.097*		SE3	0.121	0.117
Slovakia	*SK*	*0.080*	*0.133*	*Average*	*SE*	*0.141*	*0.126*
Slovenia	*SI*	*0.083*	*0.171*	United Kingdom	*UK*	*0.193*	*0.209*

Source: EU-SILC, waves 1-6 (2004-2009) cross-sectional files. Eurostat.

From the descriptive statistics, we found that 18% of male and 19% of female wage-earners in the whole sample were overqualified when we pooled all observations over the sample period. Some countries register a much lower level of overeducation, namely, Bulgaria, France, Slovakia, Slovenia and Norway. The highest overeducation rates are found in Spain, Portugal, Greece, Ireland and Lithuania. Women tend to exhibit higher levels of overeducation than men in almost every country. Exceptions are France, Greece, Portugal and Sweden. The difference between males and females overeducation rates are particularly striking in Slovenia, Hungary and Poland (Table 1). The key advantage of the EU-SILC data is the wide range of explanatory variables that we can include in our models. These can be grouped under the headings labour demand/supply, worker/job characteristics, labour market risk and institutional factors. We discuss the rationale for, and measurement of, each of these components in turn.

Labour Market Demand / Supply Indicators: Both the Job Competition Model and Assignment Theory stress the importance of the distribution of jobs relative to the stock of labour. Our models include a number of variables that measure the extent to which the supply of educated labour is outweighing demand in any country or region. Within our data we measure the stock of excess educated labour supply as (a) the ratio of third level (ISCED[8] 5[9]) graduates to employment in professional or managerial positions and (b) the rate of unemployment among ISCED 5 graduates. Previous evidence has dealt with similar sets of variables. For instance, Ghignoni and Verashchagina (2014) deployed several demand factors

---

[8] International Standard Classification of Education.
[9] Tertiary level or above.

measured at NUTS2 level (incidence of patents application, gross fixed capital formation, expenditure per worker in R&D, proportion of professionals among employed and youth unemployment). Our approach is slightly different as we explicitly focus on factors that directly address supply and demand imbalances or mismatches, whereas, with the exception of the youth unemployment rates at NUTS 2 level, Ghignoni et al. (2014) analyse demand and supply factors separately. Croce et al. (2012) and Verhaest et al. (2013) also found that measures related to excess supply of skilled labour were related to the incidence of overeducation.

Worker / Job characteristics: Both the assignment and job competition models of the labour market suggest that the distribution of both jobs and workers are potentially important in explaining the incidence of mismatch. At an aggregate level this suggests that, in addition to those controls reflecting the interaction of labour demand and supply, variables capturing the key characteristics of both labour demand (the distribution of jobs) and labour supply (the distribution of workers) may play an important explanatory role. On the labour demand side we include measures for the proportions employed in public administration, where wage rates may be influenced by non-market factors (Christofides and Michael, 2013), and in low-waged occupations, specifically in Sales & Hotels and in micro firms. To the extent to which overeducation is driven by individuals substituting higher pay for more flexible working conditions (McGuinness & Sloane, 2011) the public administration sector is included on the basis that flexible working opportunities tend to be more available there. The Sales and Hotels sector is associated with service and elementary occupations with lower entry conditions and a higher reliance on such industries, at the cost of more value added activities,

may also result in a higher incidence of overeducation, With respect to workforce supply characteristics, we include measures of (i) enrolment rates for university level programs within the region and (ii) the proportion of 25-34 year olds within a labour market on the basis that overeducation may be more of a problem in areas with a higher concentration of younger workers (Pouliakas, 2013). We also include the proportions of part-time workers and those with fixed-term contracts in the workforce. Finally, on the grounds that migrant workers are more likely to be overeducated (Battu and Sloane (2002)), we include a control for the percentage of migrants within each region.

Labour Market Uncertainty / Risk: It is highly likely that overeducation rates are more substantial in countries / regions with higher rates of returns as these are likely to generate increases in educated labour supply which, in turn, may create a temporary disequilibrium within the labour market that results in overeducation. There will certainly exist a lagged effect between any increase in rates of return and the emergence of overeducation. While increases in rates of return are indicative of a high demand for educated labour, overeducation may also emerge in response to increased educational participation if the composition of educational supply is poorly aligned with the distribution of jobs. However, not only is the length of the lag unknown but the data at hand does not enable us to generate any lagged values of the average rate of return to education. While we could include some measures of current rates of return, concerns relating to potential endogeneity preclude us from using contemporaneous variables. Nevertheless, within our framework we approximate the scale of educational returns with a measure of dispersion in rates of return on the grounds that there

will be a positive relationship between educational investment and risk. In support of the risk-return hypothesis, Pereira and Martins (2002) demonstrated, using micro 1995 data for 16 countries, a positive relationship between average rates of return and dispersion in rates of return which, they argue, is consistent with the view that rates of return to education are higher in riskier labour markets[10]. Consequently, there are grounds to support a positive relationship between overeducation and labour market risk on the grounds that risk will be associated with higher returns which could, in turn, stimulate enrolments. Our approach to risk measurement follows that of Pereira and Martins (2002) who estimate quantile regressions (QR) and proxy risk as the difference in the return to a year of schooling between the first and ninth quantiles.

Returns to schooling are obtained by estimating a standard Mincer regression for each region (or country) for each year where S_{ij} represents the years of schooling undertaken, *ex* relates to labour market experience and X_i is a vector of earnings related personal or job characteristics (equation (1)). Equation (1)[11] is estimated at both the first and ninth quantiles of the wage distribution with labour market risk approximated by subtracting the first quantile λ from that of the ninth.

$$LnW_i = \sum_{i=1}^{i} \alpha + \lambda S_i + \beta X_i + \delta ex_i + \gamma ex_i^2 + \varepsilon_i \qquad (1)$$

---

[10] In their paper the Pereira and Martins (2002) draw analogies between investments in education and the predictions of the capital asset pricing model developed by Markowitz (1952) to test the hypothesis that there exists a positive relationship between rates of return to education and the risk associated with the investment.

[11] Both the males' and females' specific equations also contained a selection term to adjust for the effects of truncation within the samples due to inactivity and/or unemployment.

The quantile regression model can be formally written as follows (see Buchinsky, 1994)

$$\ln w_i = x_i \beta_\phi + u_{\phi i} \qquad \text{with} \qquad Quant_\phi\left(\ln w_i \mid x_i\right) = x_i \beta_\phi \qquad (2)$$

where the first term on the right hand side again represents our wage equation, $Quant_\phi\left(\ln w_i \mid x_i\right)$ denotes the ϕ_{th} conditional quantile of w given x. The ϕ_{th} regression quantile, $0<\phi<1$, is defined as the solution to the problem

$$\min \beta \varepsilon R^k \left(\sum_{i:y_i \geq x_i \beta} \phi \mid \ln w_i - x_i \beta_\phi \mid + \sum_{i:y_i < x_i \beta} (1-\phi) \mid \ln w_i - x_i \beta_\phi \mid \right) \qquad (3)$$

The above equation is usually written as

$$\min \beta \varepsilon R^k \sum_i \rho_\phi\left(\ln w_i - x_i \beta_\phi\right) \qquad (4)$$

where $\rho_{\Phi(e)}$ is the check function defined as $\rho_{\Phi(e)} = \phi e$ if $\varepsilon \geq 0$ or $\rho_{\phi(e)} = (\phi-1)e$ if $\varepsilon < 0$.

It should be noted that the median estimator of $\phi = 0.5$ is a special case of the quantile regression method. The method is most usefully thought of as providing a parsimonious way of describing the wage distribution and as such it has the potential to add significantly to any empirical analysis should the relationship between the regressors and the exogenous variables evolve across the conditional wage distribution.

Institutional factors: To account for the role of institutions, in our models we control for levels of trade-union density and employment protection legislation (EPL). Finally, our models also control for the level of financial support given to students within a country or region. A priori, one might expect that overeducation would be lower in countries with lower levels of student support as the flow of qualified individuals onto the labour market at any one time will be reduced. Moreover, in most developed economies the introduction of tuition charges and income contingent loans, has represented the principal means by which governments have funded rapid expansion in the tertiary sector, which suggests that such policies may actually be positively associated with overeducation. However, if higher levels of public funding for education are linked with more meritocratic educational outcomes, then this may be also associated with better matching between skills and jobs, leading to lower rates of overeducation. While we could not find a reliable series of tuition fees data for our sample, we have included information on the level of financial aid to students in order to shed light on such factors[12].

While the majority of our variables are derived using the EU-SILC data and could thus be aggregated at regional level, this was not the case for the data on employment protection legislation (provided by OECD), trade union density (TUD)[13] and student financial aid (taken from Eurostat) both of which are available only at national level (see data appendix for details). It should be noted that

---

[12] There are, of course, many additional educational related variables that could also be of relevance to the study such as the degree of vocationalism within education provision, concentration of fields of study etc. However, such data was not readily available within EU-SILC.

[13] Trade union density figures were sourced from both the OECD and AIAS.

analysis of this type is heavily constrained by the availability of macro type indicators and small sample sizes and we believe that we have made the maximum use of the limited information available to us.

Table A2 in the data appendix shows the mean and standard deviations of the dependant and all explanatory variables in the models (with the exception of year and country dummies) broken down by gender. There are several gender related differences in some variables that may be of consequence within the multivariate environment. Namely, women register a higher share of part-time and temporary workers and are more likely to be employed in sales and hotels and in low-skilled occupations. Furthermore, the ratio of ISCED-5 graduates to employment in professional or managerial positions is much higher for women than for men suggesting that the over-supply of educated labour maybe more of a factor within female labour markets.

2.2. Methods

Given the fractional nature of our dependant variable, we estimate the fractional logit model developed by Papke and Wooldridge (1996) on the grounds that it overcomes many of the flaws that arise when Tobit and OLS models are applied to such data (Wagner (2001)). In particular, the conditional expectation of y given the explanatory variables is estimated directly and consistently, furthermore, no special adjustments are required for extreme values of the dependent variable (Papke and Wooldridge (1996)). Papke and Wooldridge (1996) propose a non-linear function for estimating the expected values of dependent variables y_i conditional on a vector of covariates x_i

$E(y_i|X_i) = G(x_i\beta)$ (5)

where G is the cumulative distribution function and β denotes the true population parameters. They chose a logistic distribution

$E(y_i | x_i) = \exp(x_i\beta) / [1 + \exp(x_i \beta)]$ (6)

and suggest the use of the Bernoulli log-likelihood function

$l_i(\beta) = Y_i \log[G(x_i \beta)] + (1-y_i) \log[1-G(x_i\beta)]$ (7)

to obtain the quasi-maximum likelihood estimator, $\hat{\beta}$. We estimate models to include country level fixed effects and dummy variables for the year the survey was conducted. In order to ensure that our estimates were not affected by the impacts of colinearity, we adopt a forward stepwise approach to ensure that the coefficients remain stable and, therefore, represent independent marginal effects. The models are estimated separately for both males and females.

3. Results

The results from our analysis are reported in tables 2 and 3. The models are presented separately for males and females. We adopt a forward stepwise specification to ensure the stability of our models and guard against the impacts of colinearity. The results reveal a variety of significant, and stable, effects that vary

somewhat by gender; however, a number of important factors are consistent across both models. The results very clearly support the view that overeducation is predominantly driven by an excess in the supply of educated labour with the ratio between the number of graduates in employment and the share of workers in professional occupations highly significant in both equations. This is consistent with job competition and assignment approaches and with previous empirical findings (Verhaest and van der Velden, 2013; Ghignoni et al., 2014). However, the marginal effects are small with the models suggesting that a 10 % increase in this ratio will drive male and female overeducation rates up by 0.5 and 0.7 parentage points respectively. Similarly, both models indicate that overeducation is highest in regions / countries where higher proportions of the 20-24 age group are enrolled in tertiary education. A ten percentage point increase in the enrolment rate is sufficient to raise the incidence of overeducation by approximately 7 percentage points for males and nearly 6 percentage points for females. However, if increased enrolments also reduce university entry standards, then the average ability levels of new graduates is likely to fall over time. Therefore, rising enrolments may increase overskilling[14] at a slower rate than overeducation on the basis that the growth in skilled graduates will be lower than the growth in total graduates. For both males and females, institutional factors such as trade-union density and the existence of employment protection legislation were found to lower the incidence of overeducation, however, the impacts varied somewhat according to gender. The EPL result, which is negative throughout the female models, and in the full specification of the male models, suggests that countries with stronger EPL are

---

[14] This describes the situation where a worker possesses skills and abilities in excess of what is required in a given job.

characterised by lower levels of overeducation. This might be because employers may take greater care in achieving good matches at recruitment because of higher costs of separation, an interpretation that is consistent with Gangl's (2004) finding in his comparison of Germany and the US. The negative effect in the female models also provides support for the assertions of McGuinness and Sloane (2010) who argue that overeducation is partially a consequence of workers decisions to occupationally downgrade in order to achieve an improved work-life balance. As labour markets with strong institutions tend to be characterized by legislation and agreements that facilitate a balance between home and family life, this will tend to reduce the extent of occupational downgrading and, hence, overeducation.

Table 2. Fractional Logit Models of Overeducation: Males (marginal effects).

	Specif. 1	Specif. 2	Specif. 3	Specif. 4	Specif. 5	Specif. 6
% foreign born in active population	0.109***	0.109***	0.103**	0.121***	0.122***	0.129***
	(0.041)	(0.042)	(0.043)	(0.042)	(0.041)	(0.040)
% part-time workers	-0.316	-0.319	-0.317	-0.257	-0.271	-0.320
	(0.183)	(0.183)	(0.187)	(0.186)	(0.196)	(0.197)
% temporary workers	-0.020	-0.020	0.021	0.042	0.054	0.065
	(0.051)	(0.051)	(0.051)	(0.050)	(0.050)	(0.048)
% employed in micro-firms	-0.011	-0.011	0.007	0.016	0.020	0.028
	(0.037)	(0.037)	(0.038)	(0.036)	(0.035)	(0.034)
% in Public Administration	-0.210*	-0.204	-0.163	-0.068	-0.042	-0.069
	(0.115)	(0.115)	(0.112)	(0.110)	(0.107)	(0.107)
% employed in Sales and Hotels	-0.072	-0.070	-0.038	-0.016	-0.030	-0.033
	(0.082)	(0.081)	(0.087)	(0.074)	(0.069)	(0.066)
% 25-34 year-olds in labour force	-0.208	-0.203	-0.177	-0.070	-0.027	-0.077
	(0.135)	(0.135)	(0.130)	(0.130)	(0.132)	(0.129)
% work and study	-0.420**	-0.423**	-0.373*	-0.413**	-0.414**	-0.340
	(0.208)	(0.208)	(0.193)	(0.194)	(0.187)	(0.184)
% employed in low skilled occupations	-0.109	-0.113	-0.095	-0.128	-0.164	-0.169
	(0.109)	(0.109)	(0.112)	(0.108)	(0.105)	(0.100)
dif_91 (country level)	0.315	0.323	0.277	0.317	0.335	0.365
	(0.237)	(0.237)	(0.231)	(0.227)	(0.237)	(0.237)
Trade Union Density	-0.040**	-0.034	-0.040**	-0.037**	-0.039**	-0.033
	(0.019)	(0.019)	(0.020)	(0.019)	(0.018)	(0.018)
Employment Protection Legislation		-0.019	-0.016	-0.022	-0.041**	-0.050**
		(0.020)	(0.019)	(0.021)	(0.017)	(0.020)
ISCED5 supply/demand ratio			0.055***	0.056***	0.055***	0.054***
			(0.014)	(0.014)	(0.013)	(0.013)
ISCED5 unemployment rate				-0.268***	-0.290***	-0.281***
				(0.096)	(0.090)	(0.087)
ISCED5&6 enrolment rates 2024					0.710***	0.762***
					(0.220)	(0.228)
financial aid to university students						-0.401***
						(0.118)
Observations	295	295	295	295	295	295
Xmfx_y	0.163	0.163	0.163	0.163	0.163	0.163
Bic	-1443	-1437	-1432	-1426	-1420	-1415
Ll	-91.22	-91.22	-91.17	-91.14	-91.12	-91.10
Chi2	7950	7809	7857	7092	7089	6754

** indicates significance at 95% and *** at 99%. ; Standard errors in parenthesis. The models include controls for both sample year and country level fixed effects which are not reported for the sake of brevity.

Table 3. Fractional Logit Models of Overeducation: Females (marginal effects)

	Specif. 1	Specif. 2	Specif. 3	Specif. 4	Specif. 5	Specif. 6
% foreign born in active population	0.031	0.036	0.042	0.044	0.034	0.035
	(0.089)	(0.089)	(0.080)	(0.080)	(0.079)	(0.079)
% part-time workers	-0.036	-0.037	-0.037	-0.018	-0.047	-0.047
	(0.087)	(0.085)	(0.073)	(0.076)	(0.074)	(0.074)
% temporary workers	0.035	0.043	0.085	0.053	0.045	0.045
	(0.049)	(0.049)	(0.049)	(0.054)	(0.054)	(0.054)
% employed in micro-firms	0.026	0.022	0.053	0.044	0.058	0.059
	(0.039)	(0.039)	(0.039)	(0.040)	(0.040)	(0.040)
% employed in Public Admin.	0.094	0.100	0.122	0.104	0.102	0.102
	(0.112)	(0.111)	(0.099)	(0.102)	(0.098)	(0.098)
% employed in Sales and Hotels	0.039	0.040	0.11*	0.12*	0.113	0.113
	(0.069)	(0.068)	(0.066)	(0.065)	(0.062)	(0.062)
% 25-34 year-olds in labour force	0.033	0.011	-0.002	-0.042	-0.043	-0.044
	(0.113)	(0.115)	(0.104)	(0.112)	(0.105)	(0.106)
% work and study	0.061	0.095	0.099	0.081	0.016	0.018
	(0.205)	(0.201)	(0.176)	(0.176)	(0.176)	(0.175)
% employed in low skilled occupations	-0.011	-0.029	-0.009	0.001	0.017	0.019
	(0.096)	(0.097)	(0.089)	(0.087)	(0.086)	(0.087)
dif_91 (country level)	0.223	0.224	0.224	0.238	0.178	0.180
	(0.254)	(0.252)	(0.233)	(0.226)	(0.226)	(0.226)
Trade Union Density	-0.021	0.005	-0.027	-0.026	-0.035	-0.034
	(0.029)	(0.027)	(0.026)	(0.026)	(0.025)	(0.025)
Employment Protection Legislation		-0.08***	-0.07***	-0.06***	-0.09***	-0.09***
		(0.022)	(0.021)	(0.021)	(0.024)	(0.025)
ISCED5 supply/demand ratio			0.071***	0.073***	0.073***	0.073***
			(0.014)	(0.013)	(0.013)	(0.013)
ISCED5 unemployment rate				0.116	0.089	0.088
				(0.093)	(0.089)	(0.089)
ISCED5&6 enrolment rates 2024					0.583***	0.583***
					(0.165)	(0.165)
financial aid to university students						-0.032
						(0.138)
Observations	295	295	295	295	295	295
Xmfx_y	0.182	0.182	0.182	0.182	0.182	0.182
Bic	-1442	-1437	-1431	-1425	-1420	-1414
Ll	-96.78	-96.76	-96.63	-96.62	-96.59	-96.59
Chi2	2451	2865	3122	3265	2783	2806

** indicates significance at 95% and *** at 99%. ; Standard errors in parenthesis.

The models include controls for both sample year and country level fixed effects which are not reported for the sake of brevity.

The negative sign of the trade-union coefficient for males would suggest that collective institutions do not increase overeducation by distorting the wage mechanism. Exactly how trade-unions reduce overeducation is unclear, one potential explanation is that the unions place upward pressure on entry job qualification as a means of preserving wages. The results are relatively unique in the context of the overeducation and certainly provide grounds for further research. However, it is important to note that the effect becomes insignificant for males when EPL is introduced into the model, suggesting that the trade-union impact may merely have been acting as a proxy control for EPL.

Within the male labour market we found that overeducation was inversely related to the graduate unemployment rate, which suggests that, perhaps not surprisingly, overeducation is lower in areas where unmatched graduates are moved into unemployment as opposed to those where they remain employed and presumably, displace workers with lower levels of schooling. Interestingly, overeducation was lower in those countries / regions where a higher proportion of employees were engaged in some form of study. This finding, although it holds true only for males, would tend to provide some support for the predictions of mobility theory, which suggests that overeducated workers are strategically entering jobs for which they are over-qualified in order to augment their educational based training with essential on-the-job training. Alternatively, the effect may be capturing the influence of vocational / apprenticeship training and or the existing of a dual system that more explicitly combines work and study, although we are unable to observe this directly in our data. Overeducation was positively related to the

migrant share of employment in the male models, a finding that is consistent with previous research (e.g. Chiswick and Miller, 2009). Overeducation among immigrants could be due to low transferability in non-domestically acquired qualifications and/ or to discrimination; however, without more detail on the nature of migrant qualifications, the relative strength of both potential explanations could not be explored further. There was only limited evidence to support the notion that overeducation was more common in riskier labour markets and, to the extent that risk is correlated with returns, that overeducation is more common in regions / countries with higher returns to education. The risk variable registered a positive coefficient in previous versions of the models but became non-significant once it was computed from wage equations including a Heckman control for sample selection bias[15]. Overeducation was inversely related to levels of financial support for male students and, while this may appear somewhat counterintuitive and contradict accepted theory[16], this effect may reflect a tendency for student financial aid to reduce inequalities in access to higher education, and thus to promote better matching between skills and jobs.

---

[15] Those factors which simultaneously explain labour market participation and wages – such as expected return to education - are also likely to induce demand for education and, as a consequence, may be correlated with the supply factors (ISCED 5 supply/demand ratio and ISCED 5&6 enrolment rates) which are shown to be strong determinants of overeducation rates.

[16] Financial support will tend to reduce the opportunity cost of education thus increasing participation and, ultimately, graduate labour supply. However, financial aid may also associated with the removal of free education (such as in the UK where income contingent loans are used to fund tuition costs) and, thus, the variable may actually by more reflective of the rising cost of education (which would reduce participation and, ultimately, overeducation). More research is needed if we are to fully understand this effect.

4. Conclusions

This research represents a unique attempt to assess, in a rigorous fashion, the extent to which overeducation is related to structural factors at a macroeconomic level. We find strong evidence to support the notion that overeducation is higher in regions where the level of educated labour supply exceeds demand and where university enrolment levels are greatest, a finding that is consistent with job competition and assignment approaches to job-skills matching, that provides confirmation that overeducation is not simply a result of the expansion of higher education. There is little evidence to support the view that overeducation is more prevalent in riskier labour markets or, to the extent that returns are correlated with risk, that overeducation is driven by high returns to education. The finding that overeducation is explained by the supply of persons with education relative to the distribution of skilled jobs in the economy with factors such as migration status and, to some extent, the sectorial composition of employment also performing an allocative role is wholly consistent with an assignment interpretation of the labour market. There is little to support the view that overeducation is consistent with short-run disequilibria that can be explained within the standard human capital framework. Furthermore, the fact that structural factors have been found to play a significant role at all again weakens assertions that overeducation represent temporary phenomena that can be attributed to either job matching models or theories of career mobility. However, the correlation between overeducation and a higher incidence of work and study in some countries may be viewed as partial evidence supporting mobility theory within the context of the male labour market. In summary, the analysis carried out using macro variables is very much in keeping with the hypothesis testing that has been conducted using micro datasets,

which generally tend to support an assignment interpretation of the labour market. Finally, the role of institutional factors represents a new finding within the literature and provides a basis for further research exploring the reasons why overeducation might be more of a problem within flexible labour markets.

The findings have a number of implications for policy. Firstly, the research suggests that considerations related to the level and nature of labour market demand in a country, or region, should form an integral part of educational planning process in terms of both the scale and composition of third-level places to be offered. The research adds to the substantial body of micro-level studies that contest the view that labour markets react automatically to changes in the level of educated labour supply, suggesting that workers can remain mismatched in employment for extended periods. Secondly, the research also points to the importance of labour market institutions in preventing overeducation within countries and regions; EPL mechanisms proved to be of more importance, particularly for females. The results provide support for the view that female overeducation is partially a consequence of workers decisions, forced or otherwise, to occupationally downgrade in order to achieve an improved work-life balance. As labour markets with strong institutions also tend to be characterized by legislation and agreements that facilitate a balance between home and family life, the result suggests that the strengthening of such policies will tend to reduce the extent of occupational downgrading and, hence, female overeducation.

5. References

Adalet McGowan, M. and Andrews, D. (2015). *Skill Mismatch and Public Policy in OECD Countries*, OECD Economics Department Working Papers, No. 1210, OECD Publishing, Paris.

Alba-Ramirez, A. (1993). Mismatch in the Spanish labor market? *Journal of Human Resources,* **28**, 259-278.

Bárcena-Martín, E., Budria, S., and Moro-Egido, A. (2012). Skill Mismatches and Wages among European University Graduates. *Applied Economics Letters,* **19**, 1471-1475.   doi: 10.1080/13504851.2011.633886.

Battu, H., Belfield, C.R., and Sloane, P.J. (1999). Overeducation among graduates: A cohort view. *Education Economics,* **7**, 21-38. doi: 10.1080/09645299900000002.

Battu, H. and Sloane, P.J. (2002). To what Extent are Ethnic Minorities in Britain Overeducated?, *International Journal of Manpower,* **23**, 192-208. doi: http://dx.doi.org/10.1108/01437720210432194

Becker, G.S. (1964). *Human Capital*. New York: Columbia University Press.

Buchinsky, M. (1994). Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression, *Econometrica,* **62**, 405–458.

Chiswick, B. R., and Miller, P. W. (2009). The international transferability of immigrants' human capital. *Economics of Education Review*, **28**, 162-169. doi:10.1016/j.econedurev.2008.07.002.

Christofides, L., and Michael, M., (2013). Exploring the public-private sector wage gap in European countries.*IZA Journal of European Labor Studies,* **32**:15, doi: DOI: 10.1186/2193-9012-2-15.

Croce G., and Ghignoni, E. (2012). Demand and Supply of Skilled Labour and

      Overeducation in Europe: A Country-level Analysis. *Comparative Economic*

      Studies, **54**, 413-419. doi:10.1057/ces.2012.12.

Di Pietro, G. (2002). Technological change, labor markets and 'low-skill, low-

      technology traps'. *Technological Forecasting and Social Change,* **69**, 885-895.

Dolton, P.J. and Vignoles, A. (2000). The Incidence and Effects of Overeducation in the

      UK Graduate Labour Market, *Economics of Education Review,* **19**, 179-198.

Gangl, M., (2004). Institutions and the Structure of Labour Market Matching in the

      United States and West Germany. *European Sociological Review,* **20**, 171-187.

      doi: 10.1093/esr/jch016.

Ghignoni, E. and Verashchagina, A. (2014). Educational qualifications mismatch in

      Europe. Is it demand or supply driven? *Journal of Comparative Economics*, **42**,

      670-692. doi:10.1016/j.jce.2013.06.006.

Groot, W. and van den Brink, H.M. (2000). Overeducation in the Labour 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.

Humburg, M., de Grip, A. and van der Velden, R. (2015). Which skills protect

      graduates against a slack labour market?, *International Labour Review*,

      published on-line, doi: 10.1111/j.1564-913X.2015.00046.x.

Jovanovic, B. (1970). Job Matching and the Theory of Turnover, *Journal of Political*

      Economy, **87**, 972-990.

Kedir, A., Kyrizi, A., and Martinez-Mora, F. (2012), *Signalling and Productivity Effects of Overeducation: Is It Really a Waste of Resources?*, University of Leicester, Dep of Economics WP, Working Paper No. 12/19, September 2012.

Mavromaras, K., McGuinness, S. and Fok, Y. (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.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, **7**, 77-91.

McGuinness, S. (2003). University Quality and Labour Market Outcomes, *Applied Economics,* **35**, 1943-1955. DOI: 10.1080/0003684032000158442.

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 Sloane, P.J. (2011). Labour Market Mismatch Among UK Graduates: An Analysis Using Reflex Data, *Economics of Education Review,* **30**, 130–145. doi:10.1016/j.econedurev.2010.07.006.

Mincer, J.A. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.

Papke, L.E. and Wooldridge, J.M. (1996). Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rate's, *Journal of Applied Econometrics*, **11**, 619–632.

Peiró, J.M, Agut, S. and Grau, R. (2010). The Relationship between Overeducation and Job Satisfaction among Young Spanish Workers: The Role of Salary, Contract of Employment and Work Experience. *Journal of Applied Social Psychology,* **40**, 666-689. doi: 10.1111/j.1559-1816.2010.00592.x.

Pereira, P.T. and Martins, P.S. (2002). Is there a return-risk link in education?, *Economics Letters,* **75**, 31-37.

Pouliakas, K. (2013). "The Skill Mismatch Challenge in Europe", chapter 6 in

    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 the existing literature*.

    OECD Social, Employment and Migration working papers No.121.

Rosen, S. (1972). Learning and Experience in the Labour Market, *Journal of Human

    Resources*, **7**, 326-342.

Sattinger, M. (1993). Assignment Models of the Distribution of Earnings, *Journal of

    Economic Literature*, **31**, 831–880.

Sicherman, N. and Galor, O. (1990). A Theory of Career Mobility. *Journal of Political

    Economy*, **9**, 101-122.

Sloane, P.J. (2003). Much Ado about Nothing? What Does the Over-education

    Literature Really Tell Us?, In Buchel, F., de Grip, A. and Mertens, A. (Eds.),

    Over-education in Europe: Current Issues in Theory and Policy, Cheltenham:

    Edward Elgar, pp. 11-45.

Thurow, L.C. (1975). *Generating Inequality*. New York: Basic Books.

Tsang, M. (1987). The Impact of Underutilization of Education on Productivity: A Case

    Study of the US Bell Companies. *Economics of Education Review,* **6**, 239-254

Tsang, M. C., Rumberger, R. W., Lewin, H. M. (1991): The Impact of Surplus

    Schooling on Worker Productivity, *Industrial Relations: A Journal of Economy

    and Society* **30**, 209-228.

Verhaest, D. and Van der Velden, R. (2013). Cross-country differences in graduate

    overeducation. *European Sociological Review*, **29**, 642-653. doi:

    10.1093/esr/jcs044.

Wagner, J. (2001). A Note on the Firm Size – Export Relationship, *Small Business Economics*, **17**, 229–237. doi: 10.1023/A:1012202405889.

Appendix A. *Definition of variables and sources[17]*

1. Overeducation rate: derived separately according to year, gender and region / country. Refers to the specific share of workers reporting a higher level of education than the mode in their two digit occupation (ISCO-88 (COM)). Levels of education are taken as ISCED.

2. Dif91 (overall): difference between coefficient of years of education in the 9th decile and the relevant coefficient of the first decile of the log of hourly gross annual/ yearly wages[18]. ISCED, region-, year- and gender–specific Mincer equations estimated with quintile regressions. Such Mincer equations control for years of education, potential experience in the labour market (and squared potential experience), nine dummies for occupations (ISCO 88) and 14 dummies for activity (NACE). In addition, selection bias is corrected à la Heckman for both gender-specific subsamples.

3. **Trade Union Density:** this variable is country and year specific. It is measured by the ratio of wage and salary earners that are trade union members, divided by the total number of wage and salary earners (OECD Labour Force Statistics). Density is calculated using survey data, wherever possible, and administrative data adjusted for non-active and self-employed members otherwise. The data were retrieved from OECD statistics data-base. The OECD series covers up to 2008, and in 2009 we have copied the 2008 data. In five countries, namely Slovenia, Latvia, Lithuania, and Estonia, we have imputed (merely linear imputation) the data for 2005 based on the information for 2000 and 2006.

4. **Employment Protection Legislation (v2):** OECD constructed series of employment protection legislation indexes. The OECD employment protection indicators are

---

[17] When no indication is made about the source it means that the source is EU-SILC and the variable has been constructed out of the micro-data set.

[18] In some countries, namely, in Spain, Greece, Italy, and Portugal only data on net income was available and, therefore, net yearly wages have been estimated in the Mincer equations.

compiled from 21 items covering three different aspects of employment protection: Individual dismissal of workers with regular contracts, additional costs for collective dismissals and regulation of temporary contracts. The data were retrieved from OECD statistics data-base. Since the OECD series covers up to 2008, in 2009 we have copied the 2008 data.

5. **Part-time rate:** This variable is year, gender and region specific. It is measured by the share of workers reporting working part-time amongst those who report being in employment.

6. **Temporality rate:** This variable is year, gender and region specific. It is measured by the share of workers reporting working in a temporary job / work contract of limited duration amongst those who report working as employees. We had to fill in the values for Denmark out of European Labour Force Survey data in Eurostat website.

7. **% of workers in micro-firms:** This variable is year, gender and region. It is measured as the specific share of workers reporting a number of workers in their local unit between one and ten, amid all types of workers.

8. **% of workers in Public Administration:** This variable is year, gender and region specific. It is measured by the share of workers reporting working in NACE "L" (Public administration and defence, compulsory social security) amid all workers. The definition was updated taking into account the recodification of NACE in 2009 (NACE – 2009).

9. **% of workers in sales and hotels:** This variable is year, gender and region specific. It is measured by the share of workers reporting working in NACE "G and H" (Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods and Hotels and restaurants) amid all workers. The definition was updated taking into account the recodification of NACE in 2009 (NACE – 2009).

10. **Share of active foreign born:** This variable is year, gender and region specific. It is measured by the share of active interviewees reporting having been born in a different country of that in which they reside.

11. **Share of 25-34 years old in the labour force:** This variable is year, gender and region specific. It is measured by the share of people aged between 25 and 34 years old amid the active population.

12. **% of workers who study:** This variable is year, gender and region specific. It is measured by the share of interviewees who are currently participating in an educational program amid those reporting being employed at the date of the interview. The person's participation in this program may be on a full-time attendance basis, a part-time attendance basis or by correspondence course. This variable only covers the regular education system (formal education, including schools, colleges and universities). In addition, if the interviewee is enrolled as an apprentice in a program within the regular education system she will be considered in education while at work as well.

13. **% of workers in low-skilled occupations:** This variable is year, gender and region specific. It is measured by the share of employed interviewees who report an occupation classified as ISCO-88 (COM) in "elementary occupations" (namely, sales and services elementary occupations, Agricultural, fishery and related laborers and Laborers in mining, construction, manufacturing and transport) over total number of employed interviewees.

14. **Ratio between ISCED-5 in employment and professional-directives:** This variable is year, gender and region specific. It is measured by the ratio between the share of workers with ISCED-5 educational attainment and the share of workers in professional-directive occupations i.e. ISCO groups I and II which consist of Legislators, senior officials and managers, Corporate managers, Managers of small

enterprises, Physical, mathematical and engineering science professionals, Life science and health professionals, Teaching professionals and Other professionals

15. **Unemployment rate among ISCED5 graduates:** This variable is year, gender and region specific. It is measured by the share of unemployed population amid active population with ISCED-5 educational attainment (higher education).

16. **Enrolment rates in higher education among 20-24 olds**: This variable is year, gender and region specific. It is measured by the share of young people (aged 20-24) who are undertaking higher education (ISCED-5&6) studies. Source: Eurostat.

17. **Financial aid**: year and country –specific financial aid to students as % of total public expenditure on education, at tertiary level of education (ISCED-5&6). Source: Eurostat. In 2009 we have taken the 2008 data since the relevant data were not available when the estimations were developed.

Appendix B. Sample description

Table A1. Number of observations per gender in multivariate analysis.

COUNTRY	REGION	males	females	COUNTRY	REGION	males	females
Austria	AT1	6	6	Greece	GR1	2	2
	AT2	6	6		GR2	2	2
	AT3	6	6		GR3	2	2
average	*AT*	*18*	*18*		GR4	2	2
Belgium	BE1	6	6	*Average*	*GR*	*8*	*8*
	BE2	6	6	Hungary	HU1	5	5
	BE3	6	6		HU2	5	5
average	*BE*	*18*	*18*		HU3	5	5
Bulgaria	BG3	2	2	*average*	*HU*	*15*	*15*
	BG4	2	2	Ireland	*IE0*	*6*	*6*
average	*BG*	*4*	*4*	Iceland	*IS*	*6*	*6*
Czech Republic	*CZ*	*5*	*5*	Italy	ITC	6	6
Germany	*DE*	*5*	*5*		ITD	6	6
Denmark	*DK*	*6*	*6*		ITE	6	6
Estonia	*EE*	*5*	*5*		ITF	6	6
Spain	ES1	6	6		ITG	6	6
	ES2	6	6	*average*	*IT*	*30*	*30*
	ES3	6	6	Lithuania	*LT0*	*5*	*5*
	ES4	6	6	Latvia	*LV0*	*5*	*5*
	ES5	6	6	The Netherlands	*NL*	*5*	*5*
	ES6	6	6	Norway	*NO*	*5*	*5*
	ES7	6	6	Portugal	*PT*	*6*	*6*
Average	*ES*	*42*	*42*	Poland	PL1	5	5
Finland	*FI*	*6*	*6*		PL2	5	5
France	FR1	5	5		PL3	5	5
	FR2	5	5		PL4	5	5
	FR3	5	5		PL5	5	5
	FR4	5	5		PL6	5	5
	FR5	5	5	*Average*	*PL*	*30*	*30*
	FR6	5	5	Sweden	SE0	4	4
	FR7	5	5		SE1	2	2
	FR8	5	5		SE2	2	2
Average	*FR*	*40*	*40*		SE3	2	2
Slovakia	*SK*	*5*	*5*	*Average*	*SE*	*10*	*10*
Slovenia	*SI*	*5*	*5*	United Kingdom	*UK*	*5*	*5*
				TOTAL		*295*	*295*

Source: EU-SILC, waves 1-6 (2004-2009) cross-sectional files. Eurostat.

Table A2. Mean and standard deviations of the dependent and the explanatory variables in the multivariate analysis.

	MALES		FEMALES	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
% Overqualified	0.181	0.095	0.193	0.077
% foreign born in active population	0.089	0.080	0.086	0.069
% part-time workers	0.048	0.024	0.251	0.145
% temporary workers	0.168	0.092	0.213	0.126
% employed in micro-firms	0.365	0.113	0.369	0.118
% employed in Public Administration	0.080	0.026	0.090	0.035
% employed in Sales and Hostels	0.156	0.036	0.202	0.054
% 25-34 year-olds in labour force	0.254	0.036	0.275	0.043
% work and study	0.039	0.019	0.054	0.027
% employed in low skilled occupations	0.089	0.040	0.125	0.045
dif_91 (country level)	0.035	0.013	0.025	0.017
ISCED5 supply/demand ratio	1.100	0.313	1.498	0.422
ISCED5 unemployment rate	0.043	0.026	0.062	0.041
ISCED5&6 enrolment rates 2024	0.267	0.045	0.350	0.065
Employment Protection Legislation [a]	2.420	0.457	2.420	0.457
Trade Union Density [b]	0.260	0.195	0.260	0.195
Financial aid to university students [c]	0.132	0.084	0.132	0.084
# of observations	295		295	

a. EPL information has been retrieved from OECD.

b. TUD figures were sourced from both the OECD and AIAS.

c. Financial aid, enrolment in tertiary education among 20-24 year-olds.

Source (for the rest of variables): EU-SILC, waves 1-6 (2004-2009) cross-sectional files. Eurostat.

Appendix C: Overeducation rates across occupations and average cell sizes

Table A3 – Gender and occupation-specific overeducation rates (%)

ISCO-88 (COM)	Males	Females
01 Armed forces	22.5	35.8
11 Legislators, senior officials and managers	2.0	2.9
12 Corporate managers	5.9	9.9
13 Managers of small enterprises	27.1	27.6
21 Physical, mathematical and engineering science professionals	0.0	0.0
22 Life science and health professionals	0.0	9.3
23 Teaching professionals	0.0	0.0
24 Other professionals	0.0	1.1
31 Physical and engineering science associate professionals	23.8	23.7
32 Life science and health associate professionals	10.9	11.1
33 Teaching associate professionals	19.1	17.9
34 Other associate professionals	30.5	27.9
41 Office clerks	24.1	24.4
42 Customer services clerks	26.1	17.9
51 Personal and protective services workers	25.0	23.2
52 Models, salespersons and demonstrators	22.7	20.2
61 Skilled agricultural and fishery workers	20.7	19.9
71 Extraction and building trades workers	18.7	28.0
72 Metal, machinery and related trades workers	19.5	19.6
73 Precision, handicraft, craft printing and related trades workers	25.8	35.6
74 Other craft and related trades workers	21.0	22.4
81 Stationary-plant and related operators	20.7	20.4
82 Machine operators and assemblers	19.6	23.7
83 Drivers and mobile plant operators	16.7	14.9
91 Sales and services elementary occupations	23.5	26.0
92 Agricultural, fishery and related labourers	26.7	30.0
93 Labourers in mining, construction, manufacturing and transport	26.8	25.7

Source: EU-SILC, waves 1-6 (2004-2009) cross-sectional files. Eurostat.