

Institute for Employment  
Research

The Research Institute of the  
Federal Employment Agency



# IAB-Discussion Paper

20/2018

Articles on labour market issues

## Patterns of Overeducation in Europe: The Role of Field of Study

Christina Boll  
Anja Rossen  
Andre Wolf

ISSN 2195-2663

# Patterns of Overeducation in Europe: The Role of Field of Study

Christina Boll (HWWI)

Anja Rossen (IAB)

Andre Wolf (HWWI)

Mit der Reihe „IAB-Discussion Paper“ will das Forschungsinstitut der Bundesagentur für Arbeit den Dialog mit der externen Wissenschaft intensivieren. Durch die rasche Verbreitung von Forschungsergebnissen über das Internet soll noch vor Drucklegung Kritik angeregt und Qualität gesichert werden.

The “IAB-Discussion Paper” is published by the research institute of the German Federal Employment Agency in order to intensify the dialogue with the scientific community. The prompt publication of the latest research results via the internet intends to stimulate criticism and to ensure research quality at an early stage before printing.

## Contents

Abstract.....	4
Zusammenfassung.....	4
1 Introduction.....	5
2 Literature Findings.....	6
3 Data and Measurement.....	8
4 Estimation method.....	10
5 Results.....	12
5.1 Descriptive results.....	12
5.2 Regression results.....	13
5.2.1 Impact of educational field.....	13
5.2.2 Country dummies.....	17
5.2.3 Other impact factors.....	19
5.2.4 Sensitivity analysis.....	21
6 Conclusion.....	23
7 References.....	25
Appendix.....	28

## Abstract

This study investigates the incidence of overeducation among graduate workers in 21 EU countries and its underlying factors based on the European Labor Force Survey 2016 (EU-LFS). Although controlling for a wide range of covariates, the particular interest lies in the role of fields of study for vertical educational mismatch. The study reveals country and gender differences in the impact of these factors. Compared to Social Sciences, male graduates from e.g. Education, Health and Welfare, Engineering, and ICT are less and those from e.g. Services and Natural Sciences are more at risk in a clear majority of countries. These findings hold for the majority of countries and are robust against a change of the standard education. However, countries show different gendered patterns of field-specific risks. We suggest that occupational closure, productivity signals and gender stereotypes answer for these cross-field and cross-country differentials. Moreover, country fixed effects point to relevant structural differences between national labour markets and between educational systems.

## Zusammenfassung

Die vorliegende Studie untersucht das Auftreten von Überqualifikation unter hochqualifizierten Beschäftigten in 21 EU-Ländern und ihre zugrunde liegenden Faktoren auf Grundlage der Europäischen Arbeitskräfteerhebung 2016 (EU-LFS). Obwohl für eine Vielzahl an erklärenden Variablen kontrolliert wird, liegt das besondere Interesse der Studie in der Rolle des Studienfaches. Die Ergebnisse zeigen, dass im Vergleich zu den Sozialwissenschaften männliche Absolventen der Studienfächer "Bildung, Gesundheit und Soziales", "Ingenieurwesen" und "Informatik und Kommunikation" ein geringeres und solche der Naturwissenschaften und Dienstleistungen ein höheres Risiko aufweisen überqualifiziert beschäftigt zu sein. Gleichzeitig weisen die verschiedenen Länder unterschiedliche geschlechtsspezifische Risikomuster auf, die auf relevante strukturelle Unterschiede zwischen den nationalen Arbeitsmärkten und zwischen den Bildungssystemen hindeuten.

**JEL classification:** J24, J21, J22

**Keywords:** Field of study, college major, overeducation, vertical mismatch, gender, realized matches, household context, EU countries, Labour Force Survey

# 1 Introduction

In general, the term overeducation refers to a job match in which the educational level of the worker clearly exceeds the educational requirements of the job. In the terminology of labour economics, this is often considered a vertical skill mismatch, as opposed to horizontal mismatches (workers choosing jobs with requirements outside the scope of their field of study/apprenticeship). A widespread occurrence of this phenomenon can seriously impair the competitiveness of an economy. From a macroeconomic perspective, an overeducation status of qualified workers reflects a waste of scarce human capital. From a microeconomic perspective, it can affect a worker's job satisfaction. In turn, a skill mismatch can reduce overall work motivation, expressing itself in more frequent absenteeism and higher turnover of the workforce (Tsang and Levin, 1985; Sicherman, 1991; Sloane et al., 1999). Moreover, overeducation is associated with earnings losses (e.g. Daly et al., 2000; Bauer, 2002; Boll and Leppin, 2016).

However, before being able to tackle the problem successfully, it is essential to understand the driving forces of overeducation at the individual level. In international comparison, the relevance of these driving forces might vary between countries and regions. Against this background, the aim of this paper is to identify possible determinants of overeducation for young (20-35 years) highly-educated (tertiary level) workers in EU-28 countries. We make use of the 2016 wave of the European Labour Force Survey (EU-LFS), a quarterly household sample survey that covers approximately 1.8 million individuals aged 15 years or older. This data set provides rich information on the respondent's demographic background, labour status, employment characteristics and educational attainment. It allows us to assess and compare the impact of a large variety of potential determinants, both separately for single countries and in a cross-country estimation. In doing so, our focus is on the role of a so far relatively neglected impact factor, the choice of field of study.

In this way, we make several contributions to the existing empirical literature on the determinants of overeducation. First, we include a range of new candidates for explanatory factors into our framework, including a person's field of study and household characteristics such as the presence of inactive and unemployed household members. Second, our results allow for a comprehensive country comparison of the associations between overeducation and distinct micro level characteristics within the EU area. This helps to identify differences in the seriousness of the phenomenon between countries and to develop tailor-made policy recipes.

Our findings reveal different overeducation risks for graduates from different fields. Compared to Social Sciences, male graduates from e.g. Education, Health and Welfare, Engineering, and ICT (Information and Communication Technologies) are less and those from e.g. Services and Natural Sciences are more at risk. These findings hold for the majority of countries and are robust against a change of the standard education. However, countries show different gendered patterns of field-specific risks. We suggest that occupational closure, productivity signals and gender stereotypes

answer for these cross-field and cross-country differentials. Moreover, country fixed effects point to relevant structural differences between national labour markets and educational systems.

## 2 Literature Findings

The empirical literature on this topic has come up with a wide range of findings on the influence of some individual- and job-related factors, in particular work experience (Alba-Ramirez, 1993; Groot, 1996; Sloane et al., 1999; Nielsen, 2011, Boll et al., 2016) and job tenure (Büchel and van Ham, 2003; Büchel and Battu, 2003; Groot and van den Brink, 2003; Ortiz, 2010) beyond macro level factors like job scarcity on the labour market which might advantage graduates due to lower expected training requirements on the side of employers (Thurow, 1975). Much less well-documented are factors related to household composition like the number of children and unemployed or inactive adults living in the same household.

Moreover, beyond the educational level also the educational field of a person, constitutes an important element for predicting labour market outcomes (van de Werfhorst and Kraaykamp, 2001; Hansen, 2001) and hence might also be a determinant of over-education. The fact that field of study is seldomly analysed so far is to some degree due to data limitations. Nevertheless, the difference might be substantial for several reasons. First, fields of study differ in their occupational focus. Fields like medicine or engineering with their quite narrowly defined job profiles might require more occupation-specific skills, raising the chances of graduates to find appropriate jobs in the corresponding occupational groups (Reimer et al., 2008). On the other hand, a narrower focus could also imply that graduates will have a harder time finding adequate jobs outside the limited scope of their field of study. In the same vein, a higher occupational closure related to high entry barriers and narrowly defined educational standards could protect graduates of fields like medicine, law or architecture from educational mismatch (Ortiz and Kucel, 2008). Thirdly, credentialism theories suggest that in a world where the true personal abilities are unknown, the chosen field of study can also act as an ability signal to employers. Obtaining a degree in fields like Maths, Natural Sciences or technical disciplines, which enjoy the reputation of imposing high intellectual demands on their students, could convince employers of the extraordinary talent and/or motivation of applicants (Barone and Ortiz, 2011). This could give them preferred access to positions with high skill requirements, possibly also outside the occupational groups associated with their subjects. Finally, field choice might be triggered by individual gender role orientations and social origin (Polachek, 1978; Bradley, 2000), such that field-specific labour market outcomes are not purely causal effects but to some part driven by selection into fields. More specifically, gender norms might impact decisions on family formation and marriage and via this channel impact educational choices (Chiappori et al., 2009; Attanasio and Kaufmann, 2017). On the other hand, having graduated as a female in a male dominated field could convey a negative productivity signal to employers, relative to male graduates in the same field,

resulting in higher overeducation. Hence, in countries with highly gender segregated higher education (e.g. Scandinavian countries; Carlsson, 2011), women graduating in gender-averse fields should be particularly penalized.

In estimating the role of field of study, it makes sense to distinguish between different levels of educational attainment. The training received by graduates from tertiary education is typically of a more academic nature and less focused on occupation-specific skills than vocational programs. Hence, the impact of the chosen field on the risk of overeducation is likely to differ with educational level. In the first analysis of this kind, Green and McIntosh (2007) restrict their estimation for the United Kingdom to the sub-sample of university graduates and thus the tertiary level. They make a quite detailed distinction between 12 educational fields. Among those, degrees in Physical Sciences and in Computing are estimated to lower the overeducation probability significantly relative to the reference category Business and Management Studies. The insignificance of the field Math explain the authors by the fact that school grades in Maths were included as an additional control variable, thereby diluting the measurement of the field effect. Moreover, signs of all field-related coefficients were negative, suggesting that the reference category business and management studies is associated with the highest overeducation risk.

In contrast, Ortiz and Kucel (2008) analysed a mixed sample of workers differing in educational attainment. Here, tertiary and non-tertiary degrees are distinguished by distinct dummies. Estimations were separately conducted for Germany and Spain. As a reference category, a tertiary degree in Social Sciences, Businesses and law was chosen. This category was associated with the lowest overeducation risk both in Germany and Spain, a result that is at least partially at odds with Green and McIntosh (2007). In fact, the large majority of other subject-degree combinations yielded significantly higher overeducation probabilities in both countries. The highest probability was estimated for tertiary graduates from the field Services, again in both countries. Moreover, both tertiary and non-tertiary graduates from Human Arts are exposed to a particularly high overeducation risk. In a further approach, Tarvid (2012) made use of the European Social Survey data and tested the field effect in a supranational sample comprising 30 countries, but only university graduates. Again, the most striking result is that graduates from Services exhibit a much higher overeducation probability than graduates from Business, Law and Economics. Probabilities lower than for the reference were detected for the fields Education and Health. Berlingieri and Zierahn (2014) compare the overeducation risk of graduates from Humanities/Social Sciences, Business/Law and Natural Sciences for highly educated German males. They find for most specifications that Business and Law graduates are at significantly higher risk than graduates from Natural Sciences. Finally, the most recent test we are aware of was conducted by Capsada-Munsech (2015) for Italian university graduates. She found that graduates from Sociopolitics experience the highest overeducation probability, even significantly higher than the reference category Humanities. The lowest probability was estimated for Medicine. Overall, even though investigations are rare and

comparability is limited by the different field classifications, the literature results suggest some considerable degree of heterogeneity, with students of Social Sciences, Services and Humanities being at higher risk than those in Natural and related Sciences.

### 3 Data and Measurement

We use data from the European Labour Force Survey (EU-LFS)<sup>1</sup> to identify possible determinants of overeducation. The EU-LFS covers approximately 1.8 Mio. individuals from the EU-28 countries (plus Iceland, Norway and Switzerland) aged 15 years or older<sup>2</sup> and asks the respondents for their demographic background, labour status, employment characteristics and their previous employment experience/search for person not in employment. Our analysis is based on 2016 data and is restricted to 21 EU-countries, guided by issues of data availability regarding household variables and occupation groups. Respondents are assigned to countries based on their place of work. In order to illustrate country differences in overeducation risk and its determinants, we perform estimations both for an aggregate cross-country sample with country fixed effects and for the single countries separately to allow for country-specific associations of the included covariates to the dependent variable.

In line with previous studies (Reimer et al., 2008; Smyth and Steinmetz, 2008), we restrict our sample to highly educated individuals, as the issue of overeducation is by definition most relevant for members of this group and, with a sharp increase of graduates' population shares during the last decades in OECD countries (from 23.3 % in 1995 to 43.1 % in 2016 on average), affecting more and more people (OECD 2018).<sup>3</sup> Highly educated individuals are defined as persons who have completed tertiary education. This corresponds to educational levels 6, 7 and 8 of the ISCED 2011 classification included in the dataset. Furthermore, the sample is restricted to respondents aged 20 to 34 years. This restriction is motivated by our primary interest in the impact of field of study, as field of study information is in EU-LFS merely available for this age group.

We refer to the above mentioned overeducation as a vertical inadequacy. In the literature, different ways for measuring overeducation are followed, from expert evaluation of occupation-specific required education (which is seldom available, Eckaus, 1964) and respondents' subjective assessments to statistical approaches (realized

---

<sup>1</sup> For more detailed information on the European Labour Force Survey, see, for example, European Union (2014).

<sup>2</sup> Norway and Sweden only cover persons between 15 and 74 years and Iceland and Switzerland only provide data on people aged 15 and more.

<sup>3</sup> Studies that compare educational groups stress the higher magnitude of overeducation for graduates. For Germany for example, a study based on the Socio-economic Panel (SOEP) estimated a 30 % (41 %) risk for West (East) German male graduates and a 36 % (38 %) risk for West (East) German female graduates of statistical overeducation in 2011 whereas the corresponding figures for workers with medium education are 8 % (12 %) for men and 14 % (10 %) for women (Boll et al., 2016).

matches). For our purposes, we adopt the variant of the realized matches approach. This is the only measure that can be employed based on the data at hand but referring to the literature, each measure has its pros and cons. Empirical evidence suggests that self-assessed overeducation is subject to other job features such as occupational status and particularly income (Dolton and Vignoles, 2000). Survey participants may be inclined to exaggerate educational requirements of their job for various reasons (Borghans and de Grip, 2000). Furthermore, self-assessed overeducation might be gender biased (Leuven and Oosterbeek, 2011).<sup>4</sup>

More specifically, we follow the realized matches approach proposed by Kiker et al. (1997) and code a person as being overeducated if his or her highest educational attainment level is higher than the modal qualification level of her occupation group at the two-digit level. We apply the 80<sup>th</sup> percentile of the levels of education within each occupation group as proposed by Ortiz and Kucel (2008). It considers a worker to be overeducated in her given job match if her educational level exceeds the 80<sup>th</sup> percentile of the distribution of observed levels of education in the given occupation. As a sensitivity check, we additionally report results calculated based on the mode as the educational standard. The choice of reference point can potentially have a sensitive impact on the measurement, depending on the specific distributions of educational levels within an occupation group. Referring the 80<sup>th</sup> percentile over the mode follows the idea that the mode regularly relates to higher overeducation rates in the same methodological setting and based on the same data (see for a literature overview Cedefop 2010, p. 18-20). This particularly applies when the underlying distribution of the dependent variable is fairly even; in this case, depending on the exact position of the most frequent single value the observations above (or below) this threshold may cover a quite high population share.

To investigate the association between overeducation risk and educational field, we implement the field of study indicator provided in the EU-LFS data as an explanatory variable. Following the classification scheme ISCED 2013-F, it distinguishes 11 field categories. In our estimation model, the single categories are coded as categorical dummy variables, choosing the category “Social Sciences, Journalism and Information” to be the omitted reference category. Additionally, in order to illuminate potential gender differences in the role of the single fields, we include interaction terms of the field category with a dummy measuring sex (female=1; male=0).

---

<sup>4</sup> The mean value is not an available option since the data provides us with educational categories only. Hence, the main advantage of the mode over the mean that it is less sensitive to outliers (Kiker et al., 1997) does not apply in our data context.

Moreover, to control for the impact of other factors, we add a broad range of control variables. In particular, we differentiate between three categories of covariates, namely personal characteristics, household characteristics<sup>5</sup> and job characteristics.

Personal characteristics include sex, marital status and two dummy variables that are equal to one if the respondent is a foreigner from another EU country or a non-EU country, respectively. Furthermore, we consider the impact of age within our already narrow sample of 20-34 years old by introducing dummies controlling for the impact of membership in the following age groups: 25-29 years old and 30-34 years old. Hence, the 20-24 years old represent the reference category in this regard. As household characteristics, we are principally able to control for the educational level (ISCED) of the spouse, the existence of unemployed or inactive adults, elderly persons (aged 75 and over) and children by age of the youngest one (between 0 and 5 years, between 6 and 11 years and between 12 and 17 years) in the same household. As recent research still identifies a strong gender bias in the labour market implications of children (Waldfogel, 1998), interaction terms with sex are included for this last indicator. However, not all of these variables are used as controls for overeducation simultaneously, as some of them are instead utilized as identification variables in a preceding Heckman correction for employment selection (see Estimation method).

Job characteristics include, among others, usual working hours and tenure. Usual working hours are given as the number of hours that a respondent is usually working per week in her main job. Tenure is defined as the number of years since a person started to work with her current employer or as self-employed. In order to shed light on potential non-linearities, both variables are also included as squared terms. Further job characteristics are considered by means of dummies that are equal to one if the respondent holds a temporary contract or if she has a second job, respectively. Firm size is split into three dummy variables, namely 11 to 19 employees, 20 to 49 employees and more than 50 employees. Persons who work for firms whose number of employees varies between 1 and 10 belong to the reference group.

As a variable reflecting the spatial dimension, the degree of urbanization at the workplace is included. It is split according to population density into two dummy variables for rural areas and towns/suburbs, with cities as the most densely-populated area as a reference category. Finally, as common control variables, we also include sector (sections according to NACE Rev. 2) and country dummies in our regressions.

## 4 Estimation method

The most simple (and also most common) approach to analyse impact factors on overeducation risk is to implement a Probit model (see Judge et al., 1988). The target

---

<sup>5</sup> Household characteristics are not available for the Nordic countries (see European Union, 2014).

variable  $y_i$  classifies a respondent either to be overeducated ( $y_i = 1$ ) or not ( $y_i = 0$ ). In the Probit model, the probability of  $y_i = 1$  is modelled as follows:

$$p = \Pr(y_i = 1|X) = \Phi(X\beta)$$

where  $\Phi(\cdot)$  is the cumulative distribution function of the standard normal distribution and  $X$  is the set of covariates presented above. The model can be estimated with the Maximum-Likelihood-Method, which yields consistent, asymptotically efficient and asymptotically normal distributed estimates. Due to the nonlinearity of the model, marginal effects are not simply given by the estimated coefficients ( $\hat{\beta}$ ), but depend on the level of the covariates.

A drawback of this simple approach is that it neglects a potential estimation bias due to self-selection into employment. It rests the analysis purely on those individuals having a job at the time of observation. However, intuition suggests that overeducation risk could well be correlated with employment selection, for instance if the prospect of entering into a skill mismatch induces job seekers to rather stay unemployed to circumvent expected earnings drawbacks or other disadvantages like job dissatisfaction. Under such circumstances, employed and non-employed individuals systematically differ in their risk levels. Results based on estimations not accounting for the impact of work selection will then be biased. Fortunately, Heckman (1979) has developed a two-step correction mechanism to take care of this issue. As a first step, based on a sample of workers and non-workers, a Probit model is specified estimating the likelihood of being in employment at the time of observation as a function of several control variables. As a second step, based on a sample restricted to workers, the Probit model for overeducation can be estimated, including the inverse mills ratio computed from the results of the first step as an additional control variable, reflecting the impact of selection to the overeducation equation.

In principle, the application of the Heckman procedure sets two data requirements. First, both workers and non-workers need to be included in the dataset. This is the case with EU-LFS. Second, in order to create exogenous variation in the inverse mills ratio, one or more identification variable(s) need to be included. These are indicators which influence the employment probability, but not directly the probability of overeducation. They therefore only appear in the employment equation. In our case, the choice of identification variables is further complicated by country heterogeneity: due to diversity of culture and traditions, different sets of identifiers could be appropriate for different countries. For a consistent approach, we applied the following specification scheme for each country sample (as well as the cross-country estimation): First, a basic Probit model for overeducation probability without selection correction including all indicators as explanatory variables was estimated. Those household variables for which the null hypothesis of zero influence could not be denied at 10 percent significance level were considered as candidates for identification variables. Second, a

Probit model for employment probability, likewise considering the whole set of indicators as explanatory variables, was estimated. Based on the results, we selected among the candidates those as identification variables for which a significant influence on employment probability could be measured. Finally, the Probit estimation of overeducation probability was repeated, this time including the sample correction (i.e. the inverse mills ratio obtained from the previous step), but omitting the identification variables. The coefficients obtained from this last regression are reported as the final results and are interpreted in the following section.

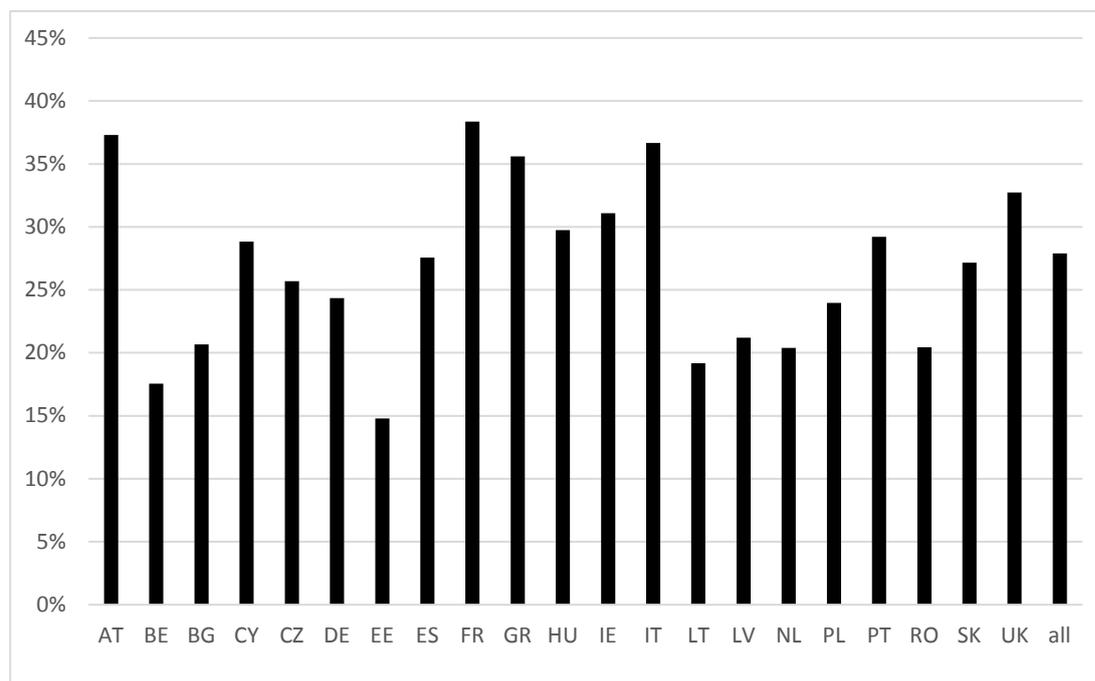
## 5 Results

### 5.1 Descriptive results

Figure 1 depicts the share of overeducated workers among workers with tertiary education at country level obtained by applying the 80<sup>th</sup> percentile as the standard education to the 2016 wave of EU-LFS. France is the country with the highest overeducation frequency, followed by Austria, Italy and Greece, while Estonia stands out with a particularly low rate. While about every second worker is considered overeducated in total, country variation is enormous. The highest rates (almost two-thirds) are measured for Italy, Ireland and the Netherlands. The lowest rates are observed for Poland, Portugal and Romania. In these countries, only about one third of the workers are assessed to be overeducated. This high degree of heterogeneity suggests that either in the single countries different impact factors are at work or that workers from different countries systematically differ in relevant characteristics or both. This can be examined based on the results of our econometric analysis discussed in the next section.

Our expectation was that using the mode instead of the 80<sup>th</sup> percentile should relate to a comparatively higher magnitude of overeducation. This is confirmed by the descriptive overeducation frequencies graphed in Figure 2 in the Appendix. About every second worker is considered overeducated in the pooled sample when the mode is used as the educational standard. Moreover, the distribution among countries shows quite a different picture. Italy, Ireland and the Netherlands represent the countries with the highest overeducation rates now, whereas the lowest rates are observed for Poland, Portugal and Romania. In these countries, only about one third of the workers are assessed to be overeducated.

**Figure 1**  
**Percentage of overeducated workers on all highly-educated workers in EU-LFS 2016, based on the 80<sup>th</sup> percentile as standard education**



Source: EU-LFS (2016).

## 5.2 Regression results

### 5.2.1 Impact of educational field

Table 1 presents estimation results for educational fields obtained in the overeducation regression at the cross-country level. Sign and significance of the single coefficients need to be interpreted relative to the reference category “male graduates from Social Sciences, Journalism and Information”, respectively.<sup>6</sup> Table 2 in the Appendix reports the country-specific results. For a correct interpretation, it is important to be aware that the 80<sup>th</sup> percentile measure represents a more restrictive criterion under most circumstances. Hence, persons classified as overeducated by this criterion can be considered severely overqualified. In what follows, we discuss the results of the cross-country estimation together with country-specific results.

We start with the base term which represents the field impact for male graduates. Compared to the reference category, graduates from Education, ICT, Engineering, and Health and Welfare exhibit a significant lower risk of overeducation whereas graduating in Natural Sciences and Services is associated to a higher risk in the cross-country comparison. Arts and Humanities, Business and Law as well as Agriculture are not significantly different from Social Sciences in terms of overeducation risk in the cross-country perspective. At country-level however, the picture differs in some aspects. For Education, the overall picture is confirmed, with the majority of countries

<sup>6</sup> We abstain from reporting results for the category of General programs as this applies only to a very small share of graduates in the tertiary segment.

reporting a comparatively lower risk (significant results: 3 with positive, 9 with negative sign). The same holds true for ICT, Health and Welfare, and Engineering, with a lower risk of overeducation compared to Social Sciences in most countries (significant coefficients signs ICT: 0 positive, 7 negative; Health and Welfare: 2 positive, 4 negative; Engineering: 1 positive, 5 negative). As in the cross-country estimation, Natural Sciences (6 positive, 0 negative) and Services (3 positive, 1 negative) are related to a higher risk compared to Social Sciences also on the country level. Moreover, Arts and Humanities which do not significantly deviate from the reference category in the country-pooled estimation, turn out to be associated with a lower risk in the majority of countries (2 positive, 4 negative). The same applies to Business and Law (3 positive, 5 negative), but the opposite holds for Agriculture which exhibits a comparatively higher risk in most countries (4 positive, 1 negative).<sup>7</sup> Concerning the significance of results, the roles of Education, ICT, Engineering, and Natural Sciences mark the most clear-cut results.

**Table 1**  
**Regression results based on the 80<sup>th</sup> percentile as the standard education**

Overeducation measure: 80 <sup>th</sup> percentile	<u>Cross-country sample</u>			
	Base term		Interaction with sex (Female=1)	
	coef	se	coef	se
<b>Sex</b>	-0.016	0.202		
<b>Field of study</b>				
General programmes	-0.170	0.401	0.341	0.488
Education	-0.219*	0.115	-0.032	0.129
Arts and humanities	-0.162	0.114	0.240*	0.131
Business and law	-0.079	0.084	0.141	0.101
Natural sciences	0.345***	0.115	-0.252*	0.145
ICT	-0.359***	0.103	0.039	0.190
Engineering	-0.196**	0.083	0.149	0.113
Agriculture	0.041	0.150	0,008	0.192
Health and welfare	-0.364***	0.119	0.089	0.133
Services	0.215**	0.108	-0.290**	0.143
Observations	34,627			

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level;

Reference category: Social Sciences, Journalism and Information.

Source: EU-LFS (2016).

The obtained relationships between field of study and vertical educational mismatch broadly fit intuition: The identified low-risk fields are for the most part associated with comparatively specific job profiles, whose well-defined qualification requirements represent entry barriers that provide preferential access for those who exhibit the appro-

<sup>7</sup> We also checked whether the picture of dominating field-effects across countries is driven by weakly significant single effects. It turned out that the reported direction of the field-effect that dominates the country-specific results applies also when all of the coefficients, including the ones which lack significance, are taken into account.

priate degree (with physicians in Health and Welfare being the most rigorous example). In this regard, Arts and Humanities seems to represent an exemption. However, beside traditionally less job-specific programs like history and philosophy, this ISCED-F category also includes more labour-market oriented subjects like handicrafts and design studies, possibly explaining the surprisingly low overeducation risk of the main group. By contrast, programs in the fields of Social Sciences are traditionally much less job-specific, forcing their graduates to compete with a range of applicants with other educational backgrounds when entering the labour market. Hence, it is no surprise to see a comparatively large share of them to be drawn into mismatches. The particularly high risk detected for graduates from Services is also as expected, given that it includes areas like Catering, Travel and Personal Services in which competition by non-academic applicants is tough. More surprising is the positive coefficient on Natural Sciences. It might be in so far explicable as this field not only contains disciplines with the reputation of setting high demands regarding analytical skills like Mathematics and Physics, but also several forms of Environmental studies, whose marketability tends to be more limited.

Confronting our results with those of the literature, we notice some parallels and discrepancies to other studies based on EU-LFS. This foremost concerns Ortiz and Kucel (2008) as well as Ghignoni and Verashchagina (2014), the only other studies we are aware of that investigate the impact of study choice on overeducation with the help of EU-LFS. The analysis of Ortiz and Kucel (2008) is based on an older version of ISCED-F with slightly different categorization. This is already visible in a distinctly composed reference category: Social Sciences, Business and Law. In their analysis, tertiary graduates from the reference category represent the group with the lowest overeducation risk. Humanities and Arts stand out with particularly high overeducation risk in both Spain and Germany, but even the fields Engineering and Health/Welfare exhibit a significantly higher risk than the reference field. This contrasts with our results for Germany where graduates from Humanities and Arts benefit from a (insignificantly) lower risk than Social Scientists do. However, our results for Spain are at least not contradictory to those of Ortiz and Kucel (2008), with an insignificantly higher risk for Spanish graduates from Arts/Humanities. The more recent study of Ghignoni and Verashchagina (2014) also comes to results comparable to Ortiz and Kucel (2008) for Germany and Spain, especially concerning the detected high risk for graduates from Engineering, compared to Services. The result for Germany is contradictory to ours, but the result for Spain accords our findings, indicating (insignificantly) lower risks for Spanish graduates from both Engineering and Services compared to Social Sciences. Concerning the other countries included in their analysis, their estimates are broadly in line with ours.

Tarvid (2012) also makes comparisons among European countries, but based on another dataset, the European Social Survey (ESS). His results are quite similar to ours. He finds a particularly high risk for graduates from Services, and low risks for graduates from Education and Health. Green and McIntosh (2007) with their analysis for

the UK differ from our and the aforementioned studies in the sense that they base their overeducation measure on workers' self-assessment and make a more detailed distinction of fields. Some parallels to our results are nevertheless noteworthy, especially concerning the low risks measured for workers with degrees in Computing and in Arts. The results of Berlingieri and Zierahn (2014) and Capsada-Munsech (2015) are of limited comparability to our analysis, because they differ too much in the way they delimit and aggregate fields. Nevertheless, some degree of consistency with our results can be observed by the facts that Berlingieri and Zierahn (2014) predict a lower overeducation risk for Engineers than for graduates from Business and Social Sciences. As our result for Business is less clear-cut as that for Engineering, their results are in line with ours. Similarly, Capsada-Munsech (2015) are in accordance with our results in the sense that they also obtain a low overeducation risk for Engineers, albeit compared to another reference category (Humanities).

The remainder of relevant studies is interested in other forms of labour market outcomes. Some of them seem to underpin the composition of our results. For instance, Nunez and Livanos (2010) analyze the impact of field choice on unemployment risk. For their Europe-wide sample, they identify the lowest risks for graduates from the fields Engineering, Education and Health/Welfare, groups that are also associated with a particularly low overeducation risk in our cross-country and the majority of country-specific estimates in our study. In a country-specific analysis, Reimer et al. (2008) likewise detect a consistently low unemployment probability for graduates from Health/Welfare, and, with just a few exceptions, also for Education and Engineering. Low overeducation and low unemployment rates both indicate a high labour market demand for the respective field-specific skills.

A slight variation of this pattern is observed when focusing on the coefficients of the interaction terms. To a large part, they are insignificant. At the same time, the base term for sex is also insignificant. Together, this implies that no difference in the overeducation risk of male and female graduates within the corresponding fields can be statistically proven. The two exceptions are *Engineering* and *Arts and Humanities*, where female graduates are assessed to be at significantly higher risk than male graduates. For a correct interpretation of this, it is important to stress again that the coefficient estimates are obtained from a regression including a large set of household- and job-related control variables. Hence, well-known reasons for gender biases on the labour market like the facts that women face tougher restrictions by the presence of children, stuck more often with part-time jobs and sort into different sectors than

men cannot serve as explanations in this case.<sup>8</sup> Instead, one could think of five (non-exclusive) channels. First, the risk discrepancies might reflect that, on average, female graduates exhibit different preferences in terms of job attributes than their male counterparts in the same educational fields. Women might prefer family-friendly and flexible work arrangements over an optimal match with corresponding higher earnings (Coudin et al., 2018). Second, they might be a sign of field-specific gender discrimination concerning access to adequate jobs, e.g. as a consequence of gender stereotypes regarding job images (Glick et al., 1995). Third, they could indicate that in these fields male graduates showed on average the better academic performance, giving them better chances to enter adequate positions. Fourth, they could also indicate the existence of educational sorting at a lower aggregation level than measured. This would mean that within these two fields female students tend to self-select into specific programs that offer comparatively worse job opportunities and are thus more prone to cause overeducation (see the discussion of educational branches in Section 5.2.2). Given the rather high aggregation level of our field variable in EU-LFS, this appears to be a likely explanation. Fifth, finally, gender differences in field-specific risks could origin in masked gender differences regarding assumed occupations.<sup>9</sup> Concerning our specific results, intuition suggests that the second explanation might be more relevant in the case of Engineering, while the fourth and fifth one seems rather suitable for Arts and Humanities.

Furthermore, gender differences in field-specific overeducation rates could origin in gender-different field-specific enrollment rates and correspondingly different demand/supply ratios on the labour market. This is not the place to discuss these effects in detail, but a reference shall be made to some empirics that analyzes the role of institutional factors underlying the observed gender segregation in the fields of study as family policies, prevalent gender norms, gender pay gaps etc. in a cross-country comparison (Zuazu, 2018; Smyth and Steinmetz, 2008).

## 5.2.2 Country dummies

The country dummies report country-specific risks that cannot be explained by the controlled individual characteristics of the national sample members (Table 3 in the Appendix). Interestingly, Latvia (Ireland, Cyprus, Spain, Slovakia) exhibit a magnitude of overeducation in the descriptive analysis (see Figure 1) which are clearly below (above) that of Germany (as a reference in the multivariate analysis) although the country dummies are insignificant. This means that the low (high) overall magnitude

---

<sup>8</sup> Strictly speaking, one potential explanatory factor we are not able to control for with our dataset is past employment. An impact of potential child-related differences in work experience between male and female graduates can thus not be excluded. However, as our sample is restricted to young workers up to 34 years old, we consider this only a very minor influence, as almost all children of the individuals in the dataset should be under aged, implying that a corresponding experience differential is captured by the interaction terms of children and sex.

<sup>9</sup> Note that we although we specify overeducation as an occupation-specific risk (see Section 1), this does naturally not prevent genders from sorting into different occupations within the same educational field.

of Latvia (Ireland, Cyprus, Spain, Slovakia) compared to Germany is fully explained by individual characteristics or, (dis-)advantageous individual portfolios are perfectly balanced with advantageous (disadvantageous) country-level effects. Poland exhibits a significantly negative country dummy although the magnitude of overeducation in Poland is quite similar to the German case, meaning that some macro-level factors in Poland outweigh the effects exerted by individual characteristics in the Polish subsample. By contrast, the significantly positive (negative) parameters of the country dummies of Austria, France, Hungary, Portugal and the United Kingdom (Belgium, Bulgaria, Estonia, Lithuania, the Netherlands and Romania) clearly indicate that beyond individual factors, some meta factors on the country level hold responsible for their comparatively higher (lower) magnitude of overeducation.

The country-specific effects may refer to country differences in Higher Education (HE) attainment rates, the skill structure of national labour markets but also to special features of educational systems, i.e. regarding selectivity of entry, drop-out rates, and the reputation of different branches of HE (masters vs. bachelors in sequential systems and universities vs. vocational schools in binary systems). According to Barone and Ortiz (2011), comparatively low attainment rates in HE in the Czech Republic, Italy, Austria and Germany should relate to relatively low overeducation rates in these countries. This holds true for the Czech Republic in our study which exerts an insignificant country dummy (compared to Germany as a reference) and displays an only slightly higher overeducation rate than Germany in the descriptive analysis. Low tertiary attainment rates and a highly stratified HE system in the Czech Republic (OECD, 2006) should play out in terms of low overeducation rates of graduates. However, Czech graduates from Agriculture suffer a significantly higher risk than Social Scientists which does not apply to Germany. However, the Czech Republic is in a more advantageous position related to Germany in terms of overall overeducation magnitude than Spain. Spain turns out as a country with high vertical mismatch which might be explained by mass enrollment in a sequential HE system generating particularly high numbers of bachelors without exhibiting a suitable absorption capacity of the high-skilled on the labour market (Barone and Ortiz, 2011). Second, in the highly segmented Spanish labour market with a high share of temporary jobs, a suboptimal match is the 'prize' for a permanent job (Ortiz, 2010). By contrast, the Dutch labour market including posts in the public sector accommodates the high supply of graduates leaving the HE system. In line with this, the Netherlands is the only West European country whose country-level factors operate more strongly against overeducation than in the German case.

With Austria and Italy however, two countries exhibit clearly higher magnitudes of overeducation than Germany and their country-level effects seem to contribute to this result. This is astonishing since Austria's HE system is highly stratified (OECD, 2006). One reason might be that vocational schools which have a shorter tradition than in Italy still send out a negative productivity signal (compared to universities), despite

posing high entry barriers. Secondly, single fields of study might drive the overall result in Austria and also in Italy. Compared to Austrian and Italian ones, German graduates from Education and Health/Welfare are at significantly lower risk than graduates from Social Sciences. This view is supported by Barone and Ortiz (2011) who state that Education and Health are among the employment areas that drive cross-country differences in overeducation.

### 5.2.3 Other impact factors

Table 4 in the Appendix presents the estimation results for the coefficients of the remaining explanatory variables at cross-country level.<sup>10</sup> In this case, the identification variable for the selection equation is eldercare, which is thus not a factor directly influencing the overeducation probability.

First, we notice the insignificance of the inverse mills ratio, implying that no association between employment selection and overeducation risk can be identified at the European level. This result is overwhelmingly confirmed by the country regressions, with Estonia and France with both positive and Slovakia with a negative significant coefficient as the only exceptions (regarding France, the significance level is 10% only).

Among the individual characteristics, only nationality is estimated to be of statistical influence at the European level. Other factors being equal, foreigners are at higher risk than domestic citizens, with non-EU foreigners suffering an even (slightly) higher risk. To the extent that foreigners include immigrants, this is in line with general economic reasoning. It would predict a higher risk for immigrants due to the non-transferability of human capital accumulated abroad and the role of cultural and language barriers, also explaining the extra risk for non-EU foreigners. However, although this result is mostly confirmed on the country level, there are some exceptions where nationality, independent of the region of origin of foreigners, does not contribute to explaining overeducation (Czech Republic, Latvia, Poland, France and the Netherlands). Low observation numbers of foreigners might account for this. Being female is not related to a higher overeducation risk per se, the same holds for being married (although on the country level, 10 (8) countries show significant associations for gender (being married), but in both directions). Moreover, a higher educational level of the spouse proves to reduce the overeducation risk in the European sample. In light of assortative mating, this is an interesting result. Even though persons with a better educated partner might tend to be more qualified themselves, they are, all things being equal, less prone to overeducation. Given our control for employment selection,

---

<sup>10</sup> Results for these impact factors at country level are available upon request.

this cannot be explained by a potential impact of the spouse's earnings on the inclination to work at all. On the country level, most countries confirm the aggregate result of this covariate, with the only exception of Greece.

Regarding the presence of children, we identify no influence on overeducation patterns for male or female individuals alike. This marks a departure from literature results such as Sloane et al. (1999), who detect a risk-enhancing effect of small children. Furthermore, heterogeneity at country level is considerable, both concerning base and interaction terms. There are several countries for which children at certain age levels are measured to be risk-enhancing for the one or the other parent or both, while they are risk-reducing in others. Results regarding the agglomeration degree are mixed on the country level, whereas the results are insignificant in the cross-country estimation. For example, living in a rural area is in four countries associated to a higher and in three other countries to a lower overeducation risk than living in a city. However, interestingly, negative base terms go hand in hand with positive female interaction terms and vice versa. Moreover, with the exception of Hungary and Latvia, the absolute amounts of the effects of interaction terms are in these countries larger than those of the base terms. This means that risk differences between cities and rural areas work in opposite directions for men as for women. With respect to the job variables, merely the size of the firm and the type of the working contract are determined to be of relevance at cross-country level. Persons who work in firms with 11-19 employees are *ceteris paribus* predicted to have the lowest overeducation risk, with however only a slight difference to the impact of firms with 20-49 employees. Hence, workers in very small firms have the highest risk, but workers in large firms (50 and more employees) rank second. At country level, the picture is diverse however. Latvia, Poland, Germany and Slovakia stand out with lowest risks in large firms, in Portugal and Greece this applies to the 11-19 category and in some others, firm size is not at all relevant. Furthermore, workers with temporary contracts face a higher risk than workers in permanent positions. This is both in line with expectation and the results of Green and McIntosh (2007) and Ortiz (2010). One explanation could be that the transitory nature of fixed-time jobs could convince people to accept less ideal matches. In this regard, there is no contradictive evidence on the country level although sometimes significance is lacking. More usual working hours are associated with a lower overeducation risk in the cross-sectional estimation, which is widely confirmed on the country level with only four countries showing significant opposite effects (Czech Republic, France, Ireland and Latvia) and the negative squared term indicates that the positive hours effect operates at decreasing marginal rates. A quite robust result is the decreasing effect of tenure; only Cyprus and Latvia present opposing results here, with largely positive squared terms). Furthermore, many industry effects turn out to be highly significant in the aggregate estimation, with diverse patterns on the country level. The Education sector stands out with a risk-reducing effect through-

out countries and also in the cross-country estimation. This also applies to Public Administration and Defense as the second component of the public sector, with the only exceptions of Cyprus (on the 10% level significant) and Slovakia.

#### 5.2.4 Sensitivity analysis

As we pointed out in the methodological section, our statistical approach of measuring overeducation is not the only consistent method available. Therefore, it is important to grasp the sensitivity of a measurement choice by comparing it to an alternative. To learn in how far the application of this alternative might impact our econometric results, we repeated our estimations, this time based on the mode instead as the 80<sup>th</sup> percentile. Table 5 in the Appendix lists the estimates for the fields of study obtained under this alternative scenario for the cross-country sample and Table 6 in the Appendix displays the country-specific results. Focusing on the base term first, we can state that male graduates from the disciplines Education, Arts and Humanities, ICT, Engineering and Health and Welfare all face a significantly lower overeducation risk than the reference group at cross-country level. The lowest risk is estimated for male graduates from ICT, all else being equal. Significantly higher risks than for the reference category are measured for male graduates from Natural Sciences and Services. For the remaining fields, a difference to the reference category cannot be established statistically. On the country-level, the lower risk for graduates in the field of Education compared to Social Sciences is confirmed (significant parameters Education: 3 positive<sup>11</sup>, 5 negative). The same holds true for Arts and Humanities (2 positive<sup>12</sup>, 8 negative), ICT (1 positive<sup>13</sup>, 6 negative), Business and Law (1 positive, 3 negative)<sup>14</sup>, Engineering (2 positive<sup>15</sup>, 7 negative), and Health and Welfare (3 positive<sup>16</sup>, 5 negative). By contrast, Services (6 positive, 2 negative<sup>17</sup>) and Natural Sciences (5 positive, 1 negative<sup>18</sup>) are associated with a higher overeducation risk than Social Sciences. Furthermore, also Agriculture which lacks significance in the overall estimation, is related to a comparatively higher risk (5 positive, 3 negative). In the cross-field comparison, results are most clear-cut for Services, Arts/Humanities, ICT, Engineering and

---

<sup>11</sup> For example, France and the Netherlands exhibit significantly higher overeducation probabilities for Education than for the reference category.

<sup>12</sup> Concerning Arts and Humanities, country heterogeneity is remarkably low, with France and Belgium as the only two countries in which graduates face a significantly higher overeducation probability than the reference graduates from the Social Sciences.

<sup>13</sup> This positive association applies to Greece.

<sup>14</sup> Regarding Business and Law, there are three countries for which the base term is significantly negative instead of insignificant (Cyprus, Estonia, Slovakia), while there is one (Greece) for which it is positively significant.

<sup>15</sup> For example, in Belgium overeducation probability is significantly higher for engineers than for (male) social scientists.

<sup>16</sup> Health and Welfare as the perhaps most job-focused field exhibits three country outliers (Czech Republic, Lithuania, Netherlands) in which the field-specific risk is measured to be significantly higher than for the Social Sciences.

<sup>17</sup> Outliers concerning the base term of Services are Greece and Lithuania. For Greece, this might reflect the important role of the tourism sector and its related training courses as a job motor for the young generation.

<sup>18</sup> Greece stands out with its significantly negative coefficient for Natural Sciences.

Natural Sciences. Note that the results regarding the prevalent direction (“lower” vs. “higher”) compared to the reference category does not only hold true for the significant parameters on the country level but also when the total of results is taken into account.

In comparison of measures and in terms of the basic effect, the pattern of fields derived from the mode is almost identical with the one derived from the 80<sup>th</sup> percentile, both for the cross-section and the country-specific results. That is, our sensitivity analyses confirm the main results regarding the base term. Not a single field is changing sign due to the measure change. Directly compared based on significant results, the switch from the 80<sup>th</sup> percentile to the mode is accompanied with even more clear-cut results with respect to Services and Arts/Humanities whereas results for Education, Natural Sciences, and ICT are a bit more clear-cut under the stricter overeducation threshold. Results remain similarly unequivocal for Engineering. Less clear-cut results are retrieved, irrespective of the measurement, for Agriculture, Health/Welfare and Business/Law. As a tendency, severe overeducation seems to play a lower role for Arts/Humanities (where graduates are at low risk anyway) and Services, but marks a notable portion of overeducation among engineers, teachers, ICT graduates (although their overall risk is low), and national scientists (with a quite substantial risk).

Most of the country heterogeneity regarding overeducation however concerns the interaction term (gender-specific effects) although most of them are insignificant at cross-country level. One example are female graduates from Education which are at significantly higher risk than male ones in several countries (Cyprus, Estonia, Greece, Portugal, Romania), while there are at least two in which female graduates are at significantly lower risk (Ireland, Netherlands). Concerning Arts and Humanities, the slightly higher risk of females (significant at 10 % level) at the cross-country level is also confirmed for a wide range of countries which accords with the findings based on the 80<sup>th</sup> percentile. In contrast to the cross-country results, among Natural Scientists there are three countries in which males are subject to systematically higher risk than females (Czech Republic, France, Ireland) and one in which the opposite is true (Lithuania). Concerning ICT, the positive association to overeducation that is measured for Greece applies to male workers only. At the same time, a positive association can be testified for females only in Austria and Romania. Agriculture shows a particularly high degree of heterogeneity, also in gender interactions with overeducation risks. Concerning Health and Welfare, the higher risks in the Czech Republic, Lithuania, and the Netherlands associated with overeducation, almost exclusively apply to male graduates. In Germany, Poland, Bulgaria, Estonia, Greece, Lithuania, female graduates from Engineering face higher risks than male graduates in the same field. In Germany and Poland, where male engineers exhibit lower overeducation risks than male social scientists, female engineers approach the risk level of male social scientists when the interaction and the base term of Engineering are taken together.

Concerning the country dummies, a somewhat different picture emerges, compared to those derived from the standard education in the main analysis (see Table 7 in the

Appendix). As the associations of fields of study are robust against the measure change, the deviations in country-fixed effects have to be attributed to deviating associations of individual characteristics with overeducation, resulting from the change in overeducation measurement.

Regarding other individual and household characteristics, Table 8 in the Appendix reports detailed results. The mode differs from the 80<sup>th</sup> percentile employed as the standard education in the main analyses in that under the mode, individuals living together with persons who are unemployed or inactive on the labour market face a higher overeducation risk. This was not the case under the 80<sup>th</sup> percentile as the much stricter overeducation threshold. One interpretation could be that the need to financially support unemployed household members induces workers to avoid own unemployment by accepting some degree of suboptimal matches but this does not extend to very bad matches. Alternatively, it could point at a linkage between household composition and job-related productivity based on assortative mating (Mare, 1991): workers living together with unemployed might on average be less productive themselves, a fact that reduces their chances to find a match adequate to their formal education. At country level, both variables represent identification variables for a large set of countries, are thus not included in the overeducation regressions. Among those in which they are included, there are a few deviations to the cross-country result to be noticed. This primarily concerns Hungary, with opposite (i.e. positive) contributions measured for both variables.

Finally, we also undertook additional estimations including further explanatory factors at the regional level (NUTS 2), such as the regional unemployment rate and the employment-to-population ratio. However, due to the large share of missing values, models including this regional information did not yield reliable results for the population as a whole.

## 6 Conclusion

The purpose of this paper was to conduct a comprehensive econometric analysis of potential determinants of overeducation among graduates in 21 EU countries in a unified framework. A special focus was set on the role of subject choice made by the individuals during their educational career. It turned out that both in the cross-country estimation and at country level differences in overeducation risk between graduates from different fields are significant. Furthermore, gender discrepancies in the impact of certain fields are noticeable. At the European level, graduates from Services, Natural Sciences and Agriculture are found to exhibit the highest risk among men. At the same time, male graduates from fields like ICT, Health/Welfare, Education, Engineering but interestingly also Arts/Humanities, are exposed to a rather low risk. The field-specific risks apply for the majority of countries and are robust against a measure change in the educational standard. We suggest that occupational closure and productivity signals are among the relevant underpinnings of these cross-field differentials.

Gender differences in field-specific overeducation risks mostly lack statistical significance, with Engineering and Arts and Humanities, where female graduates are assessed to be at significantly higher risk than male graduates, marking the exceptions. By and large, the above named sensitivity analysis deploying another standard education confirms this pattern for an alternative method of measuring overeducation. Prevalent gender stereotypes, discrimination, but also gendered preferences and institutional factors on the national level boosting gendered segregation of college majors are supposed to drive gendered overeducation, but we have to left this issue for a more detailed analysis in future research.

Moreover, country fixed effects point to relevant structural differences between national labour markets and educational systems. As we included a selection correction in our estimation approach, country differences concerning employment selection should not be the source of this heterogeneity. Rather, differences in educational systems, in the capacities of labour markets to absorb young tertiary graduates as well as in culture- and tradition-based attitudes seem likely candidates. Although we made some references to the literature here, disentangling these different national-aspects and utilizing them for an analysis of country patterns represents a second interesting avenue for future research.

Further arguments add to the limitations of our study. Despite the wide range of individual covariates we are aware of missing factors that proved to be relevant for overeducation propensity like paternal background (Jackson et al., 2008) or students' academic ability before enrolling in higher education (Barone and Ortiz, 2011). With the underlying econometric approach, causal interferences must not be drawn. Finally, we are aware of the sensitivity of results with respect to the measure of overeducation. Results often change when subjective evaluations of overeducation are used instead of the statistical measure (Bauer, 2002; Chiswick and Miller, 2010; Nielsen, 2011; Boll et al., 2016). Therefore, thirdly, it would be interesting to evaluate our results based on different specification of the target variable. Unfortunately, with the data at hand, this was not possible.

## References

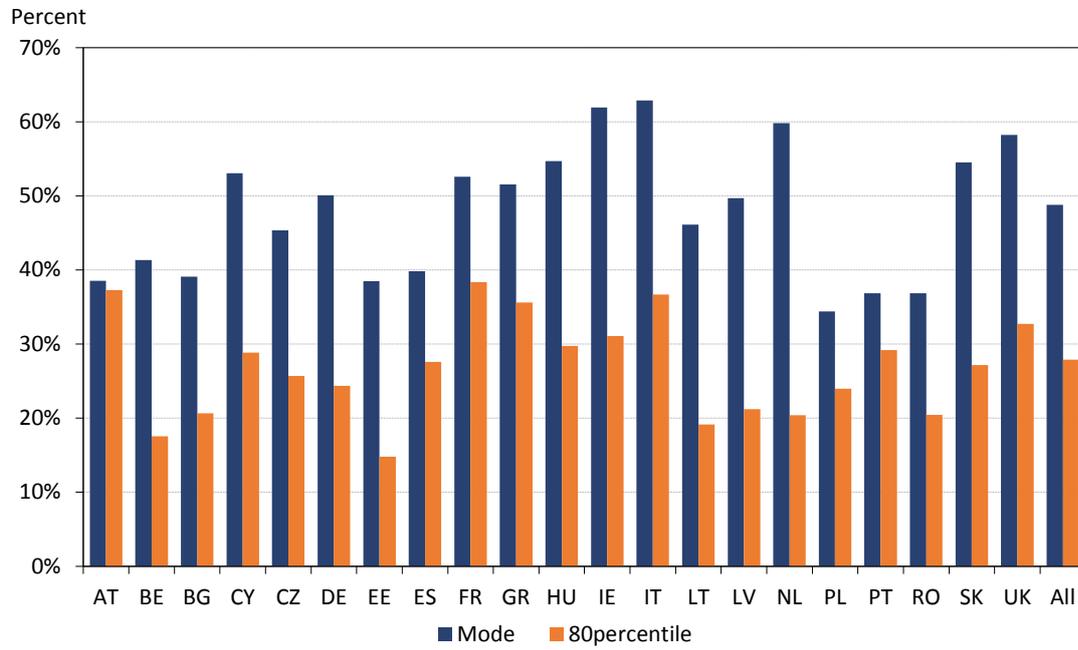
- Alba-Ramirez, A. (1993): Mismatch in the Spanish labor market: overeducation? In: *The Journal of Human Resources* 28(2), 259–278.
- Attanasio, O. P.; Kaufmann, K. M. (2017): Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations. In: *Journal of Economic Behaviour and Organization* 140, 35–55.
- Barone, C.; Ortiz, L. (2011): Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation. In: *Higher Education* 61(3), 325–337.
- Bauer, T. (2002): Educational mismatch and wages: a panel analysis. In: *Economics of Education Review* 21(3), 221–229.
- Berlingieri, F.; Zierahn, U. (2014): Field of study, qualification mismatch, and wages: Does sorting matter? (No. 14-076). ZEW-Zentrum für Europäische Wirtschaftsforschung/Center for European Economic Research.
- Boll, C.; Leppin, J. (2016): Differential Overeducation in East and West Germany: Extending Frank's Theory on Economic Returns Changes the Picture of Disadvantaged Women. In: *LABOUR: Review of Labour Economics and Industrial Relations* 30(4), 455–504.
- Boll, C.; Leppin, J.; Schömann, K. (2016): Who is overeducated and why? Probit and dynamic mixed multinomial logit analyses of vertical mismatch in East and West Germany. In: *Education Economics* 24(6), 639–662.
- Borghans L.; de Grip, A. (2000): The Debate in Economics About Skill Utilization, in Borghans L. and de Grip A. (eds.): *The Overeducated Worker?* In: *The Economics of Skill Utilization*, Cheltenham, UK: Edward Elgar: 3–23.
- Bradley, K. (2000): The Incorporation of Women into Higher Education: Paradoxical Outcomes? In: *Sociology of Education* 73(1), 1–18.
- Büchel, F.; Battu, H. (2003): The theory of differential overqualification: does it work? *Scottish Journal of Political Economy* 50(2), 1-16.
- Büchel, F.; Van Ham, M. (2003): Overeducation, regional labor markets, and spatial flexibility. In: *Journal of Urban Economics*, 53(3), 48–493.
- Capsada-Munsech, Q. (2015): The role of social origin and field of study on graduates overeducation: the case of Italy. In: *Higher Education* 69(5), 779–807.
- Carlsson, M. (2011): Does hiring discrimination cause gender segregation in the Swedish labor market? In: *Feminist Economics* 17(3), 71–102.
- Cedefop (European Centre for the Development of Vocational Training) (2010): *The skill matching challenge. Analysing skill mismatch and policy implications*, Luxembourg: Publications Office of the European Union, 2010.
- Chiappori, P.-A.; Iyigun, M.; Weiss, Y. (2009): Investment in schooling and the marriage market. In: *American Economic Review* 99(5), 1689–1713.
- Chiswick, B. R.; P. W. Miller (2010): Does the Choice of Reference Levels of Education Matter in the ORU Earnings Equation? In: *Economics of Education Review* 29(6), 1076–1085.
- Coudin, E., Maillard, S., & To, M. (2018). *Family, firms and the gender wage gap in France* (No. W18/01). Institute for Fiscal Studies.

- Daly, M. C.; Büchel, F.; Duncan, G. J. (2000): Premiums and penalties of surplus and deficit education. Evidence from the United States and Germany. In: *Economics of Education Review* 19(2), 169–178.
- Dolton P.; Vignoles A. (2000): The Incidence and Effects of Overeducation in the U.K. Graduate Labour Market. *Economics of Education* 19(2), 179–198.
- Eckaus R. (1964): Economic Criteria for Education and Training. In: *Review of Economics and Statistics* 46(2), 181–190.
- European Union (2014): Quality report of the European Union Labour Force Survey 2013–2014 edition.
- Ghignoni, E.; Verashchagina, A. (2014): Educational qualifications mismatch in Europe. Is it demand or supply driven? In: *Journal of Comparative Economics*, 42(3), 670–692.
- Glick, P.; Wilk, K.; Perreault, M. (1995): Images of occupations - Components of gender and status in occupational stereotypes. In: *Sex roles*, 32(9-10), 565–582.
- Green, F.; McIntosh, S. (2007): Is there a genuine under-utilization of skills amongst the over-qualified? In: *Applied Economics* 39(4), 427–439.
- Groot, W.; van den Brink, H. (2003): The dynamics of skill mismatches in the Dutch labour market. Overeducation in Europe. In: *Current Issues in theory and Policy*, 49–63.
- Groot, W. (1996): The incidence of, and returns to overeducation in the UK. In: *Applied Economics* 28(10), 1345–1350.
- Hansen, M.N. (2001): Education and Economic Rewards: Variations by Social-Class Origin and Income Measures. In: *European Sociological Review* 17(3), 209–31.
- Heckman, J. J. (1979): Sample selection bias as a specification error. In: *Econometrica* 47(1), 153–161.
- Jackson, M.; Lujckx, R.; Pollak, R.; Vallet, L.-A.; van de Werfhorst, H. G. (2008): Educational Fields of Study and the Intergenerational Mobility Process in Comparative Perspective. In: *International Journal of Comparative Sociology* 49(4-5), 369–388.
- Judge, G.; Hill, R.; Griffiths, W.; Lütkepohl, H.; Lee, T.-C. (1988): Introduction to the theory and practice of econometrics – 2nd edition.
- Kiker, B. F.; Santos, M. C.; De Oliveira, M. M. (1997): Overeducation and undereducation: Evidence for Portugal. In: *Economics of Education Review* 16(2), 111–125.
- Leuven, E.; Oosterbeek, H. (2011): Overeducation and Mismatch in the Labour Market. IZA Discussion Paper 5523.
- Mare, R. D. (1991): Five decades of educational assortative mating. In: *American Sociological Review*, 15–32.
- Nielsen, C.P. (2011): Immigrant over-education: evidence from Denmark. In: *Journal of Population Economics* 24(2), 499–520.
- Nunez, I.; Livanos, I. (2010): Higher education and unemployment in Europe: an analysis of the academic subject and national effects. In: *Higher Education* 59(4): 475–487.
- OECD (2018): Population with tertiary education (indicator), DOI: 10.1787/0b8f90e9-en (Accessed on 14 August 2018).

- OECD (2006): *Education at a glance 2006*. Paris: OECD. ISBN: 9264025316.
- Ortiz, L. (2010): Not the right job, but a secure one: overeducation and temporary employment in France, Italy and Spain. In: *Work, employment and society*, 24(1), 47–64.
- Ortiz, L.; Kucel, A. (2008): Do fields of study matter for over-education? The cases of Spain and Germany. In: *International Journal of Comparative Sociology* 49(4-5), 305–327.
- Polachek, S.W. (1978): Sex Differences in College Major. In: *Industrial and Labor Relations Review* 31(4), 498–508.
- Reimer, D.; Noelke, C.; Kucel, A. (2008): Labor market effects of field of study in comparative perspective: An analysis of 22 European countries. In: *International Journal of Comparative Sociology*, 49(4-5), 233–256.
- Sicherman, N. (1991): "Overeducation" in the labor market. In: *Journal of Labor Economics* 9(2), 101–122.
- Sloane, P.J.; Battu, H.; Seaman, P.T. (1999): Overeducation, undereducation and the British Labour market. In: *Applied Economics* 31(11), 1437–1453.
- Smyth, E.; Steinmetz, S. (2008): Field of study and gender segregation in European labour markets. In: *International Journal of Comparative Sociology*, 49(4-5), 257–281.
- Tarvid, A. (2012): Effects of the Field of Higher Education on the Prevalence of Over-education in European Countries. In: *Education: Evaluation, Reform and Policy*, 167–184.
- Thurow, L. (1975): *Generating inequality*. New York: Basic Books.
- Tsang, M. C.; Levin, H. M. (1985): The economics of overeducation. In: *Economics of Education Review* 4 (2), 93-104.
- van de Werfhorst, H.G.; Kraaykamp, G. (2001): Four Field-Related Educational Resources and Their Impact on Labor, Consumption, and Sociopolitical Orientation. *Sociology of Education* 74(4): 296–317.
- Waldfoegel, J. (1998): Understanding the "family gap" in pay for women with children. In: *Journal of Economic Perspectives*, 12(1), 137–156.
- Zuazu, I. (2018): *Cultural Values, Family Decisions and Gender Segregation in Higher Education: Evidence from 26 OECD Economies*, University of the Basque Country – Departamento de Fundamentos des Análisis Económico I, Working Paper Series IL. 107/18.

## Appendix

**Figure 2**  
**Distribution of overeducation rates under different measures of standard education**



Source: EU-LFS (2016).

**Table 2**  
**Country-specific regression results – Coefficients of the field-of-study dummy variables based on the 80<sup>th</sup> percentile as the standard education**

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	<b>Austria</b>				<b>Belgium</b>				<b>Bulgaria</b>			
<b>Sex</b>	-0.417	0.674			0.670*	0.397			3.141	1.931		
<b>Field of study</b>												
General programmes												
Education	-0.670*	0.358	0.512	0.388	0.390	0.260	-0.414	0.291	6.056***	1.781	-5.683***	1.758
Arts and humanities	-0.351	0.360	0.715*	0.415	0.351	0.226	0.033	0.276				
Business	-0.078	0.270	0.244	0.321	-0.319*	0.187	0.173	0.231	4.466***	0.809	-3.066***	0.975
Natural sciences	0.116	0.346	0.190	0.452	0.842***	0.255	-0.736**	0.347	4.432***	1.135	-1.964	1.674
ICT	-1.088***	0.320	1.609***	0.544	-0.720***	0.264	0.310	0.532	-0.708	1.326	2.052	1.869
Engineering	-0.507*	0.267	0.383	0.350	0.337*	0.178	0.041	0.280	4.486***	0.602	-3.367***	1.100
Agriculture	1.234**	0.485	-1.171*	0.616	0.381	0.351						
Health and welfare	-0.583	0.359	0.171	0.390	0.355	0.252	-0.732**	0.284	-0.373	0.834		
Services	-0.123	0.505	-0.190	0.640	0.186	0.324	-0.089	0.404				
Observations	1,536				2,650				124			
	<b>Cyprus</b>				<b>Czech Republic</b>				<b>Estonia</b>			
<b>Sex</b>	-2.294**	1.028			-3.268***	0.713			2.237*	1.203		
<b>Field of study</b>												
General programmes												
Education	-14.916***	0.654	15.563***	0.747	0.225	0.622	-1.246*	0.704	-4.774***	0.771	6.378***	1.125
Arts and humanities	1.579***	0.550	-1.685***	0.653	1.010	0.846	-1.270	0.929	-0.388	0.764	2.156*	1.273
Business	-0.108	0.332	0.222	0.450	0.122	0.542	-0.831	0.619	-0.692	0.502	2.297**	1.003
Natural sciences	3.227***	0.807	-3.107***	0.867	1.860***	0.710	-2.169***	0.814	0.813	0.825	1.038	1.242
ICT					-0.039	0.543						
Engineering	-0.114	0.462	0.683	0.698	0.368	0.476	-0.188	0.614	0.599	0.500	1.898*	1.048
Agriculture					1.782***	0.654	-1.364	0.851	1.357	0.849		
Health and welfare	-0.337	0.662	-5.804***	0.749	-4.400***	0.549	3.609***	0.634	0.947	0.766	1.223	1.096
Services	0.621	0.550			1.428***	0.529	-0.841	0.773	-0.472	0.584	0.472	1.130
Observations	562				449				378			

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	France				Germany				Greece			
<b>Sex</b>	-0.562	0.554			0.044	0.262			-6.713***	0.488		
<b>Field of study</b>												
General programmes					-0.355	0.585	0.779	0.719				
Education	-0.170	0.825	-0.212	0.863	-0.533***	0.183	-0.044	0.214	-1.003**	0.505	1.365***	0.528
Arts and humanities	-0.224	0.633	0.493	0.667	-0.111	0.195	0.111	0.229	-0.093	0.373	0.373	0.434
Business	-0.119	0.402	0.440	0.432	-0.067	0.155	0.059	0.181	0.444	0.292	-0.760**	0.361
Natural sciences	1.146**	0.454	-0.767	0.539	0.379**	0.178	-0.397*	0.226	-0.738	0.467	0.863	0.567
ICT	-0.068	0.439	-0.508	0.737	-0.556***	0.175	0.120	0.279	0.473	0.291	-0.571	0.421
Engineering	-0.050	0.396	-0.264	0.458	-0.350**	0.155	0.013	0.204	-0.698**	0.285	1.076***	0.357
Agriculture	0.633	0.792	-0.561	0.908	0.046	0.268	-0.562	0.348	-0.122	0.471	0.441	0.568
Health and welfare	0.046	0.482	0.490	0.524	-0.687***	0.212	0.049	0.235	0.181	0.368	0.248	0.408
Services	0.120	0.447	-0.263	0.568	0.384**	0.195	-0.268	0.248	-1.237***	0.349	0.994**	0.418
Observations	1,088				7,379				1,443			
	<b>Hungary</b>				<b>Ireland</b>				<b>Italy</b>			
<b>Sex</b>	4.456***	0.523			1.182	0.828			0.131	0.277		
<b>Field of study</b>												
General programmes					0.629	0.647						
Education	0.542	0.353	-0.352	0.378	0.024	0.596	-1.540**	0.667	-0.471	0.531	-0.182	0.559
Arts and humanities	-0.154	0.317	0.213	0.417	-4.457***	0.648	3.335***	0.733	-0.278	0.319	0.497	0.352
Business	0.393	0.279	-0.020	0.330	0.165	0.518	-0.966	0.605	0.095	0.265	-0.031	0.311
Natural sciences	0.580	0.381	-1.380**	0.536	0.784	0.546	-1.605**	0.645	0.368	0.320	-0.489	0.392
ICT	-0.600**	0.305	0.237	0.672	-0.374	0.532	-1.133	0.692	-1.510***	0.313	1.773***	0.546
Engineering	-0.231	0.247	-0.263	0.405	0.545	0.516	-1.164*	0.670	-0.027	0.241	-0.181	0.318
Agriculture	-0.101	0.345	1.208***	0.447					0.114	0.504	2.410***	0.608
Health and welfare	-6.249***	0.585	6.276***	0.593	0.398	0.583	-1.664**	0.667	0.211	0.330	-0.790**	0.354
Services	0.660	0.442	-0.719	0.524	-0.462	0.688	-0.630	0.865	0.167	0.465	0.040	0.528
Observations	1,458				939				1,943			

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	Latvia				Lithuania				Netherlands			
<b>Sex</b>	6.111***	1.303			-1.898**	0.964			0.044	0.660		
<b>Field of study</b>												
General programmes									0.230	0.821		
Education	10.278***	0.883	-9.803***	0.848	2.207**	0.858	-2.374**	0.924	0.474	0.406	-0.060	0.456
Arts and humanities	5.088***	0.908	-5.039***	0.957	-4.896***	0.861	5.236***	0.939	-0.870**	0.432	1.681***	0.504
Business	8.972***	0.559	-8.562***	0.591	-0.092	0.463	0.440	0.524	-0.549*	0.296	0.823**	0.362
Natural sciences					-0.064	0.572	0.955	0.702	0.043	0.400	0.796	0.518
ICT					0.099	0.512						
Engineering	8.128***	0.579	-6.972***	0.657	0.389	0.453	0.469	0.547	-0.641*	0.356	0.603	0.581
Agriculture	3.572***	0.746	-3.703***	1.049					-0.178	0.634	0.121	0.823
Health and welfare	4.918***	0.938	-5.628***	0.990	-3.179***	0.496	3.420***	0.553	-0.633	0.434	0.523	0.484
Services	9.686***	0.899	-10.193***	1.133	-0.108	0.537	0.745	0.667	-0.000	0.343	0.134	0.445
Observations			440				914				1,422	

	Poland				Portugal				Romania			
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	Poland				Portugal				Romania			
<b>Sex</b>	-0.046	0.386			2.409***	0.643			1.015	0.795		
<b>Field of study</b>												
General programmes												
Education	-0.056	0.161	0.355*	0.195	-0.818*	0.432	1.642***	0.490	-1.454***	0.434	2.762***	0.562
Arts and humanities	-0.230	0.197	0.398	0.244	-1.134***	0.371	1.078**	0.479	-0.011	0.541	0.174	0.571
Business	-0.220*	0.129	0.391**	0.160	-0.656*	0.349	0.831*	0.435	0.669***	0.250	-0.554*	0.297
Natural sciences	0.189	0.190	0.561**	0.231	-1.082	0.732	1.197	0.850	0.727	0.454	-0.316	0.538
ICT	-0.728***	0.155	0.075	0.311					0.050	0.354	0.295	0.560
Engineering	-0.351***	0.131	0.908***	0.184	-1.237***	0.340	1.078**	0.455	-0.264	0.272	-0.164	0.358
Agriculture	0.011	0.366	0.457	0.419	0.897	0.709	-1.107	0.815	-0.002	0.457	0.245	0.870
Health and welfare	-0.010	0.216	-0.019	0.251	-1.140***	0.414	0.929**	0.464	1.366**	0.597	-1.394**	0.662
Services	0.054	0.170	0.387*	0.219	0.186	0.363	-0.424	0.498	0.426	0.333	-0.122	0.520
Observations			5,749				1,063				1,626	

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	Slovak Republic				Spain				United Kingdom			
<b>Sex</b>	-1.577**	0.625			-0.851	0.627			0.271	0.449		
<b>Field of study</b>												
General programmes												
Education	-20.702***	1.534	20.159***	1.601	-1.474**	0.612	-0.522	0.757	-0.820*	0.466	0.407	0.518
Arts and humanities	-5.795***	0.791	6.713***	0.895	1.326	1.021	-1.660	1.121	-0.184	0.261	-0.075	0.315
Business	-0.675	0.446	1.935***	0.704	-0.983**	0.494	0.803	0.606	-0.057	0.253	-0.046	0.310
Natural sciences	1.426*	0.815	0.130	1.000	-1.046	0.688	1.443	0.905	0.054	0.274	-0.344	0.345
ICT	-0.803	0.576			-1.638***	0.613	0.112	1.065	-0.054	0.296	-0.525	0.483
Engineering	0.143	0.509	1.892**	0.764	-0.813	0.594	-0.573	0.838	0.197	0.274	-0.168	0.401
Agriculture	2.034***	0.729	-0.214	1.068	-1.477	1.004	0.672	1.290	-4.340***	0.517	4.111***	0.625
Health and welfare	0.192	0.824	-0.232	0.837	-1.238	0.786	0.615	0.847	-0.472	0.435	-0.395	0.468
Services	0.466	0.498			-0.488	0.646	-0.777	0.862	0.262	0.363	-0.457	0.475
Observations	565				417				1,462			

Reference category: Social Sciences, Journalism and Information.

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level;

Source: EU-LFS (2016).

**Table 3**  
**Regression results – Country dummies in the cross-country sample based on the 80<sup>th</sup> percentile as the standard education**

Overeducation measure: 80 <sup>th</sup> percentile	Cross-country sample	
Country dummy	coef	se
Austria	0.302***	0.049
Belgium	-0.228***	0.041
Bulgaria	-0.357***	0.114
Cyprus	0.060	0.064
Czech Republic	0.094	0.074
Estonia	-0.346***	0.092
France	0.367***	0.053
Greece	0.309***	0.052
Hungary	0.270***	0.051
Ireland	0.019	0.057
Italy	0.452***	0.044
Latvia	-0.340***	0.062
Lithuania	-0.053	0.084
Netherlands	-0.314***	0.052
Poland	-0.210***	0.037
Portugal	0.206***	0.058
Romania	-0.222***	0.052
Slovak Republic	-0.042	0.073
Spain	0.005	0.083
United Kingdom	0.208***	0.045
Observations	34,627	

Reference category: Germany

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level;

Source: EU-LFS (2016).

**Table 4**  
**Regression results – Coefficients of non-study related variables in the cross-country sample based on the 80<sup>th</sup> percentile as the standard education**

Overeducation measure: 80 <sup>th</sup> percentile	Cross-country sample			
	Base term		Interaction with sex (Female=1)	
	<i>coef</i>	<i>se</i>	<i>coef</i>	<i>se</i>
Inverse mills ratio	0.092	0.087		
<u>Household variables</u>				
Unemployed (dummy)	0.178	0.119		
Inactive (dummy)	0.035	0.060		
Education spouse	-0.012***	0.004		
<u>Age of youngest child</u>				
< 5 years	0.007	0.050	0.045	0.066
6-11 years	0.141	0.104	0.006	0.131
12-17 years	0.209	0.345	0.225	0.413
<u>Individual variables</u>				
Sex	-0.016	0.202		
Marital Status	0.053	0.051	-0.095	0.064
<u>Age group (ref: 20-24)</u>				
25-29 years	-0.177	0.171	0.057	0.194
30-34 years	-0.065	0.173	-0.115	0.195
<u>Nationality (ref: Domestic)</u>				
Foreigner EU countries	0.540***	0.062		
Foreigner non EU countries	0.5553***	0.077		
<u>Job variables</u>				
<u>Firm size (ref: &lt; 10)</u>				
11-19 persons	-0.193***	0.054		
20-49 persons	-0.191***	0.046		
50 and more persons	-0.114***	0.041		
Temporary contract	0.209***	0.039		
Working hours (in 10h)	-0.206***	0.076		
Working hours squared (in 10h)	0.007	0.011		
Tenure (in 10y)	-0.381***	0.117		
Tenure squared (in 10y)	0.154*	0.087		
<u>Degree of urbanization (ref: cities)</u>				
Towns and suburbs	0.018	0.063	0.019	0.063
Rural area	-0.045	0.074	-0.046	0.074
Observations	34,627			

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level; Dummies for nationality and industry (sections NACE Rev.2) included.

Source: EU-LFS (2016).

**Table 5**  
**Regression results based on the mode as the standard education**

Overeducation measure: Mode	Cross-country sample			
	Base term		Interaction with sex (Female=1)	
	coef	se	coef	se
<b>Sex</b>	-0.116	0.173		
<b>Field of study</b>				
General programmes	0.426	0.403	-0.801	0.500
Education	-0.236**	0.099	-0.144	0.111
Arts and humanities	-0.216**	0.105	0.228*	0.124
Business and law	0.008	0.074	0.047	0.089
Natural sciences	0.286***	0.107	-0.116	0.133
ICT	-0.306***	0.088	0.026	0.168
Engineering	-0.266***	0.076	0.247**	0.112
Agriculture	-0.108	0.156	0.062	0.196
Health and welfare	-0.237**	0.107	-0.079	0.118
Services	0.314***	0.097	-0.160	0.132
Observations	34,624			

Reference category: Social Sciences, Journalism and Information.

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level;

Source: EU-LFS (2016).

**Table 6**  
**Country-specific regression results – Coefficients of the field-of-study dummy variables based on the mode as the standard education**

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	<b>Austria</b>				<b>Belgium</b>				<b>Bulgaria</b>			
<b>Sex</b>	-0.387	0.689			0.010	0.321			0.084	1.633		
<b>Field of study</b>												
General programmes												
Education	-0.765**	0.356	0.600	0.388	0.026	0.213	-0.377	0.238	2.706	1.807	-1.408	1.934
Arts and humanities	-0.438	0.358	0.753*	0.413	0.361*	0.194	0.052	0.239	2.260	1.583	-0.450	1.788
Business	0.096	0.264	0.065	0.316	-0.062	0.157	0.131	0.193	0.687	0.964	0.564	1.130
Natural sciences	0.072	0.344	0.231	0.453	0.677***	0.229	-0.284	0.300	1.761	1.497	-0.645	1.665
ICT	-1.165***	0.318	1.746***	0.543	0.153	0.185	-0.639	0.390	0.737	1.098	1.918	1.284
Engineering	-0.521**	0.264	0.450	0.346	0.375**	0.156	0.092	0.250	0.469	0.92	2.347*	1.202
Agriculture	1.126**	0.472	-1.018*	0.606	0.466	0.312						
Health and welfare	-0.649*	0.356	0.224	0.388	0.226	0.220	-0.355	0.245	0.982	0.774		
Services	-0.190	0.500	-0.154	0.634	0.396	0.259	0.258	0.343	-1.761	1.316	3.732**	1.687
Observations	1,536				2,650				199			
	<b>Cyprus</b>				<b>Czech Republic</b>				<b>Estonia</b>			
<b>Sex</b>	-3.492***	0.895			0.568	0.840			-3.418***	0.941		
<b>Field of study</b>												
General programmes												
Education	-3.228***	0.704	3.171***	0.770	0.074	0.554	-0.699	0.612	-4.568***	0.694	4.698***	0.815
Arts and humanities	-0.474	0.624	0.480	0.681	-4.533***	0.549	4.460***	0.697	-1.052	0.740	0.687	0.931
Business	-1.019***	0.368	0.913**	0.450	0.136	0.501	-0.132	0.587	-1.063**	0.439	0.979	0.642
Natural sciences	-0.415	0.679	0.523	0.746	2.031***	0.631	-2.339***	0.743	-0.173	0.775	0.330	1.020
ICT					-0.360	0.507	-0.792	1.018	-3.025***	0.787		
Engineering	-1.460***	0.435	0.467	0.652	0.083	0.417	-0.293	0.563	-0.405	0.425	1.790***	0.689
Agriculture					0.984*	0.591	-1.100	0.788	-0.116	0.777		
Health and welfare	-0.931*	0.52	0.320	0.572	1.001**	0.496	-1.254**	0.541	-0.239	0.655	1.421*	0.791
Services	-0.062	0.533			1.622***	0.573			0.142	0.534	1.840**	0.745
Observations	642				446				405			

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	<b>France</b>				<b>Germany</b>				<b>Greece</b>			
<b>Sex</b>	-0.167	0.582			0.042	0.259			-4.741***	0.495		
<b>Field of study</b>												
General programmes					-0.144	0.536	-0.064	0.675				
Education	1.382**	0.610	-0.816	0.652	-0.745***	0.169	0.233	0.198	-1.378***	0.499	1.145**	0.518
Arts and humanities	1.062*	0.546	-0.874	0.593	-0.527***	0.186	0.548**	0.219	-0.712	0.459	0.473	0.498
Business	0.473	0.377	-0.165	0.411	0.090	0.149	0.289*	0.175	0.718*	0.371	-1.561***	0.416
Natural sciences	1.505***	0.434	-1.367***	0.506	0.461***	0.172	-0.075	0.215	-1.049**	0.515	0.875	0.581
ICT	0.330	0.405	-0.92	0.653	-0.175	0.160	0.301	0.249	0.573*	0.345	-1.030**	0.438
Engineering	0.192	0.377	-0.440	0.452	-0.385***	0.147	0.498***	0.190	-1.119***	0.308	0.672*	0.374
Agriculture	1.452*	0.824	-1.524	0.930	-0.058	0.273	-0.112	0.342	-0.567	0.541	0.560	0.627
Health and welfare	0.121	0.463	-0.137	0.505	-0.604***	0.178	0.059	0.204	-0.571	0.421	0.025	0.445
Services	0.863*	0.443	-0.932*	0.541	0.535***	0.201	0.007	0.256	-0.838***	0.314	-0.226	0.37
Observations	1,088				7,379				1,443			
	<b>Hungary</b>				<b>Ireland</b>				<b>Italy</b>			
<b>Sex</b>	0.100	0.659			1.686**	0.762			-0.159	0.313		
<b>Field of study</b>												
General programmes												
Education	-0.005	0.332	-0.344	0.357	0.287	0.57	-2.055***	0.676	-0.537	0.480	-0.122	0.511
Arts and humanities	-0.547	0.349	0.473	0.404	-0.311	0.634	-0.835	0.767	-0.749**	0.363	0.896**	0.405
Business	-0.159	0.299	0.452	0.345	0.330	0.510	-1.014	0.631	-0.230	0.303	0.424	0.355
Natural sciences	0.021	0.442	-0.155	0.512	0.498	0.543	-1.268*	0.678	0.425	0.411	-0.537	0.484
ICT	-0.954***	0.294	-0.316	0.648	0.014	0.524	-0.843	0.704	-0.186	0.441	0.610	0.597
Engineering	-0.464*	0.258	-0.391	0.358	0.415	0.507	-1.596**	0.697	-0.578**	0.271	-0.102	0.363
Agriculture	-1.032***	0.37	2.031***	0.455					-0.689	0.508	2.718***	0.605
Health and welfare	-10.118***	0.433	9.980***	0.451	0.349	0.566	-1.563**	0.681	0.012	0.359	0.036	0.386
Services	0.575	0.411	-1.089**	0.497	0.41	0.615	-1.433*	0.811	0.934*	0.533	0.243	0.628
Observations	1,445				941				1,858			

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	<b>Latvia</b>				<b>Lithuania</b>				<b>Netherlands</b>			
<b>Sex</b>	-7.713***	1.105			-1.307*	0.757			1.372**	0.592		
<b>Field of study</b>												
General programmes									-1.398	0.860		
Education	-0.377	0.697	0.671	0.752	0.836	0.679	-0.763	0.721	0.730**	0.371	-1.038**	0.407
Arts and humanities	-5.891***	0.546	6.107***	0.652	-1.373**	0.572	1.134*	0.619	-0.038	0.397	0.365	0.466
Business	-0.244	0.478	0.586	0.534	-0.385	0.295	0.808**	0.345	-0.263	0.286	-0.370	0.333
Natural sciences					-0.501	0.424	1.240**	0.538	0.528	0.38	0.236	0.525
ICT					-0.187	0.372	0.016	0.545				
Engineering	0.463	0.547	0.299	0.672	-0.054	0.283	0.804**	0.384	-0.061	0.308	0.308	0.444
Agriculture	-7.048***	0.827	7.508***	1.230					0.636	0.606	-1.077	0.784
Health and welfare	0.357	0.844	-0.745	0.889	4.289***	0.335	-3.805***	0.387	0.671*	0.349	-0.646*	0.382
Services	-1.018	0.765	0.896	0.999	-1.186***	0.415	1.579***	0.532	0.317	0.364	-0.216	0.443
Observations	430				944				1,419			
	<b>Poland</b>				<b>Portugal</b>				<b>Romania</b>			
<b>Sex</b>	0.276	0.383			2.750***	0.652			-0.439	0.632		
<b>Field of study</b>												
General programmes												
Education	-0.058	0.154	0.230	0.181	-1.235**	0.596	1.908***	0.638	-2.050***	0.404	3.267***	0.484
Arts and humanities	-0.478**	0.188	0.468**	0.224	-1.042***	0.372	1.150**	0.469	-0.540	0.419	1.033**	0.462
Business	-0.124	0.118	0.080	0.144	-0.390	0.342	0.643	0.417	0.311	0.198	-0.097	0.244
Natural sciences	0.214	0.187	0.332	0.223	-0.962	0.805	1.023	0.917	0.288	0.445	0.241	0.504
ICT	-0.666***	0.142	-0.328	0.333					-0.547*	0.316	0.872*	0.478
Engineering	-0.337***	0.122	0.698***	0.170	-1.014***	0.337	0.668	0.441	-0.686***	0.230	0.335	0.295
Agriculture	-0.021	0.338	0.102	0.393	0.735	0.730	-0.334	0.824	0.062	0.356	0.224	0.597
Health and welfare	-0.023	0.195	-0.365*	0.215	-0.726*	0.395	0.487	0.430	-0.017	0.510	-0.182	0.513
Services	0.257	0.166	0.205	0.214	0.261	0.359	0.435	0.542	0.097	0.301	0.449	0.454
Observations	5,747				1,063				1,641			

	Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)		Base term		Interaction with sex (Female=1)	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
	<b>Slovak Republic</b>				<b>Spain</b>				<b>United Kingdom</b>			
<b>Sex</b>	0.708	0.525			-0.719	0.697			-0.223	0.404		
<b>Field of study</b>												
General programmes												
Education	-0.455	0.449	0.104	0.556	-0.465	0.692	-0.225	0.826	-0.523	0.319	-0.142	0.374
Arts and humanities	-5.290***	0.454	6.523***	0.617	0.984	0.972	-1.348	1.087	-0.346	0.236	-0.016	0.294
Business	-0.616*	0.347	1.721***	0.472	-0.308	0.585	0.204	0.690	0.034	0.225	-0.265	0.289
Natural sciences	1.545**	0.657	-0.531	0.759	-0.677	0.789	1.060	1.011	0.028	0.258	-0.068	0.329
ICT	-1.401***	0.471			-0.928	0.631	0.604	0.939	-0.412	0.270	-0.588	0.450
Engineering	1.391***	0.355	-0.111	0.567	-0.043	0.657	-0.819	0.851	0.326	0.259	-0.038	0.402
Agriculture	1.023*	0.545	-0.415	0.914	-1.452	0.915	0.166	1.250	-1.069*	0.634	0.861	0.761
Health and welfare	-0.308	0.688	0.768	0.673	-0.966	0.786	0.573	0.832	-0.481	0.336	-0.364	0.372
Services	1.130**	0.489			0.335	0.716	-0.535	0.937	0.662**	0.331	-0.938**	0.446
Observations	541				417				1,448			

Reference category: Social Sciences, Journalism and Information.

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level;

Source: EU-LFS (2016).

**Table 7**  
**Regression results – Country dummies in the cross-country sample based on the mode as the standard education**

Overeducation measure: Mode	Cross-country sample	
Country dummy	coef	se
Austria	-0.429***	0.047
Belgium	-0.198***	0.036
Bulgaria	-0.707***	0.102
Cyprus	0.055	0.061
Czech Republic	-0.113*	0.067
Estonia	-0.385***	0.074
France	-0.005	0.052
Greece	-0.014	0.049
Hungary	0.168***	0.050
Ireland	0.225***	0.051
Italy	0.527***	0.045
Latvia	-0.140***	0.054
Lithuania	-0.002	0.078
Netherlands	0.218***	0.045
Poland	-0.700***	0.033
Portugal	-0.390***	0.055
Romania	-0.513***	0.045
Slovak Republic	-0.045	0.062
Spain	-0.451***	0.077
United Kingdom	0.229***	0.042
Observations	34,624	

Reference category: Germany

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level;

Source: EU-LFS (2016).

**Table 8**  
**Regression results – Coefficients of non-study related variables in the cross-country sample based on the mode as the standard education**

Overeducation measure: Mode	Cross-country sample			
	Base term		Interaction with sex (Female=1)	
	<i>coef</i>	<i>se</i>	<i>coef</i>	<i>se</i>
Inverse mills ratio	-0.029	0.071		
<u>Household variables</u>				
Unemployed (dummy)	0.270***	0.103		
Inactive (dummy)	0.087	0.053		
Education spouse	-0.013***	0.003		
<u>Age of youngest child</u>				
< 5 years	-0.065	0.044	0.094	0.059
6-11 years	-0.018	0.101	0.153	0.123
12-17 years	-0.285	0.331	0.442	0.392
<u>Individual variables</u>				
Sex	-0.116	0.173		
Marital Status	0.016	0.046	-0.032	0.058
<u>Age group (ref: 20-24)</u>				
25-29 years	-0.113	0.137	0.107	0.163
30-34 years	-0.039	0.136	-0.044	0.162
<u>Nationality (ref: Domestic)</u>				
Foreigner EU countries	0.410***	0.064		
Foreigner non EU countries	0.475***	0.074		
<u>Job variables</u>				
<u>Firm size (ref: &lt; 10):</u>				
11-19 persons	-0.066	0.053		
20-49 persons	-0.117**	0.046		
50 and more persons	-0.111**	0.041		
Temporary contract	0.162***	0.036		
Working hours (in 10h)	-0.025	0.065		
Working hours squared (in 10h)	-0.011	0.009		
Tenure (in 10y)	-0.062	0.108		
Tenure squared (in 10y)	-0.050	0.080		
<u>Degree of urbanization (ref: cities)</u>				
Towns and suburbs	-0.013	0.044	0.049	0.057
Rural area	0.059	0.053	0.070	0.068
Observations	34,624			

\*: Statistical significance at 10%-level; \*\*: Statistical significance at 5%-level; \*\*\*: Statistical significance at 1%-level; Dummies for nationality and industry (sections NACE Rev.2) included.

Source: EU-LFS (2016).

## Recently published

No.	Author(s)	Title	Date
<a href="#">6/2018</a>	Schäffler, J. Moritz, M.	German FDI in the Czech Republic – Employment effects in the home country	2/18
<a href="#">7/2018</a>	Fuchs, J. Weber, B.	Fachkräftemangel: Inländische Personalreserven als Alternative zur Zuwanderung	2/18
<a href="#">8/2018</a>	Wapler, R. Wolf, K. Wolff, J.	Do active labour market policies for welfare recipients in Germany raise their regional outflow into work?	3/18
<a href="#">9/2018</a>	Wanger, S. Zapf, I.	For better or worse? How more flexibility in working time arrangements and fatherhood affect men's working hours in Germany	3/18
<a href="#">10/2018</a>	Warning, A. Weber, E.	Digitalisation, hiring and personnel policy: evidence from a representative business survey	3/18
<a href="#">11/2018</a>	Stepanok, I.	FDI and Unemployment, a Growth Perspective	3/18
<a href="#">12/2018</a>	Knize, V.	Migrant women labor-force participation in Germany	4/18
<a href="#">13/2018</a>	Schierholz, M.; Brenner, L.; Cohausz, L.; Damminger, L.; Fast, L.; Hörig, A.; Huber, A.; Ludwig, T.; Petry, A.; Tschischka, L.	Eine Hilfsklassifikation mit Tätigkeitsbeschreibungen für Zwecke der Berufskodierung	5/18
<a href="#">14/2018</a>	Janser, M.	The greening of job in Germany	5/18
<a href="#">15/2018</a>	Dettmann, E. Weyh, A. Titze, M.	Heterogeneous effects of investment grants - Evidence from a new measurement approach	5/18
<a href="#">16/2018</a>	Speidel, M. Drechsler, J. Jolani, S.	R Package hmi: A Convenient Tool for Hierarchical Multiple Imputation and Beyond	6/18
<a href="#">17/2018</a>	Fuchs, M. Fackler, D. Hölscher, L. Schnabel, C.	Do startups provide employment opportunities for disadvantaged workers?	6/18
<a href="#">18/2018</a>	Bellmann, L. Brixy, U.	Hiring by start-ups and regional labor supply	7/18
<a href="#">19/2018</a>	Stepanok, I.	Trade and FDI, the Proximity-Concentration Trade-Off Revisited	

As per: 2018-08-30

For a full list, consult the IAB website <http://www.iab.de/de/publikationen/discussion-paper.aspx>

## Imprint

**IAB-Discussion Paper 20/2018**  
11 September 2018

### Editorial address

Institute for Employment Research  
of the Federal Employment Agency  
Regensburger Straße 104  
D-90478 Nuremberg

### Editorial staff

Ricardo Martinez Moya, Jutta Palm-Nowak

### Technical completion

Renate Martin

### All rights reserved

Reproduction and distribution in any form, also in parts,  
requires the permission of IAB Nuremberg

### Website

<https://www.iab.de>

### Download of this Discussion Paper

<http://doku.iab.de/discussionpapers/2018/dp2018.pdf>

ISSN 2195-2663

### For further inquiries contact the author:

Anja Rossen  
Phone +49 911 179-4158  
E-mail [Anja.Rossen4@iab.de](mailto:Anja.Rossen4@iab.de)