

Magnitude and Impact Factors of the Gender Pay Gap in EU Countries

Christina Boll, Julian Leppin,
Anja Rossen, André Wolf

This report was financed by, and prepared for the use of the European Commission, Directorate-General for Justice and Consumers, Unit D2 "Equality between men and women", in the framework of a contract managed by Fondazione Giacomo Brodolini (FGB) in partnership with Istituto per la Ricerca Sociale (IRS).

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this report. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.



FGB - Fondazione Giacomo Brodolini

Via Solferino 32
00185 Rome
Italy
Tel +39 064424 9625
Fax +39 0644249565
www.fondazionebrodolini.it



IRS - Istituto per la Ricerca Sociale

Via XX Settembre 24
20123 Milano
Italy
Tel. +39 2467 641
www.irs-online.it




Europe Direct is a service to help you find answers
to your questions about the European Union.

Freephone number (*):

00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.



MAGNITUDE AND IMPACT FACTORS OF THE GENDER PAY GAP IN EU COUNTRIES

Christina Boll, Julian Leppin,
Anja Rossen, André Wolf




Table of Contents

Non-technical Summary and Implications	8
Methodological Guides	20
MODULE 1. OVEREDUCATION	21
1. Introduction	21
2. Literature Findings	23
2.1 Theories	23
2.2 Empirical results	25
3. Measurement and Data	33
4. Model Set-up	36
5. Results	38
5.1 Descriptive results	38
5.2 Estimation results	43
6. Conclusion	54
7. References – Module 1	56
MODULE 2: GENDER PAY GAP – SES (2010)	60
1. Introduction	60
2. Literature	62
2.1 Theoretical motivation of the gender wage divide	62
2.2 Empirical evidence on wage gap determinants	64
3. Measurement and Data	69
4. Model	71
5. Results	73
5.1 Decomposition in explained and unexplained gender pay gap	73
5.2 Decomposition of the explained gender pay gap	76
5.3 Factors behind the unexplained gender pay gap	81
6. Conclusion	85
MODULE 3: GENDER PAY GAP – EU-SILC (2012)	87
1. Introduction	87
2. Measurement and Data	89

3. Results	92
3.1 Decomposition into explained and unexplained gap	92
3.2 Decomposition of the explained gender pay gap	95
3.3 Decomposition of the unexplained part of the gender pay gap	100
4. Conclusion	103
5. References – Modules 2 and 3	105
SUGGESTIONS FOR ADDITIONAL STATISTICAL INDICATORS	111
Appendix – Module 1	113
Appendix – Module 2	137
Appendix – Module 3	142

List of Figures

Unadjusted, explained and unexplained gender pay gap, EU average (22 countries), based on SES 2010 data, in %	9
Decomposition of the unadjusted gender pay gap in 2010, in %	11
Decomposition of the unexplained part of the gender pay gap (in %), 2010	13
Share of overeducated workers on the workforce by gender (all countries)	16
Risk maps for the overeducation probability of high- and medium-skilled workers	17
Main drivers of gender earnings divides and corresponding combating strategies	19
Figure 1: Percentage of overeducated on all medium-skilled employees, 2013	38
Figure 2: Percentage of overeducated on all high-skilled employees, 2013	39
Figure 3: Share of overeducated workers by skill level and worker origin (cross-country sample), 2013	42
Figure 4: Marginal effects of the interaction of gender and marital status (reference: unmarried men)	45
Figure 5: Marginal effects of field of study for male workers (cross-country estimation), 2013	48
Figure 6: Marginal effects of no. of children for female workers in five West European countries, 2013	50
Figure 7: Marginal effects of firm size (cross-country estimation), 2013	53
Figure 8: Relationship between gender pay gap and female employment in SES	75
Figure 9: Decomposition of the gender pay gap (in %), 2010	77
Figure 10: Decomposition of the unexplained part of the gender pay gap (in %), 2010	82
Figure 11: Size of the unadjusted gender pay gaps in EU-SILC and SES	92
Figure 12: Estimation Procedure in Module 3	93
Figure 13: Explained gender pay gap according to SES and EU-SILC	95
Figure 14: Decomposition of the gender pay gap (in %) in EU-SILC (2013)	96
Figure 15: Decomposition of the unexplained gender pay gap (in %) in EU-SILC (2013)	100

List of Figures – Appendices

Figure A 1: Relationship between gender pay gap and female employment in EU-SILC	142
--	-----

List of Tables

Unadjusted, explained and unexplained gender pay gaps in EU countries based on SES 2010 data, in %	10
Table 1: Percentage of overeducated and not overeducated persons by sex, age, and educational attainment	40
Table 2: Estimation results (cross-country estimation)	43
Table 3: Unadjusted, explained and unexplained gender pay gap based on SES 2010 data, in %	74

List of Tables - Appendices

Table A 1: Descriptive statistics: high-skilled workers, EU-28	113
Table A 2: Descriptive statistics: medium-skilled workers, EU-28	115
Table A 3: Estimation results high-skilled workers: Austria, Belgium, Bulgaria, Cyprus, and Czech Republic	117
Table A 4: Estimation results high-skilled workers: Germany, Denmark, Estonia, Spain, and Finland	119
Table A 5: Estimation results high-skilled workers: France, Greece, Croatia, Hungary, and Ireland	121
Table A 6: Estimation results high-skilled workers: Italy, Lithuania, Luxembourg, Latvia, and the Netherlands	123
Table A 7: Estimation results high-skilled workers: Portugal, Romania, Sweden, Slovakia, and the United Kingdom	125
Table A 8: Estimation results medium-skilled workers: Austria, Belgium, Bulgaria, Cyprus, and Czech Republic	127
Table A 9: Estimation results medium-skilled workers: Germany, Denmark, Estonia, Spain, and Finland	129
Table A 10: Estimation results medium-skilled workers: France, Greece, Croatia, Hungary, and Ireland	131
Table A 11: Estimation results medium-skilled workers: Italy, Lithuania, Luxembourg, Latvia and the Netherlands	133
Table A 12: Estimation results medium-skilled workers: Portugal, Romania, Sweden, Slovakia, and the United Kingdom	135
Table A 13: Composition of the explained gender pay gaps at country level (in %), 2010	137
Table A 14: Drivers of the occupation-related endowment effect (cross-country estimation) in SES (2010)	139
Table A 15: Drivers of the industry-related endowment effect (cross-country estimation)	140
Table A 16: Composition of the unexplained gender pay gaps at country level	140
Table A 17: Unadjusted, explained and unexplained gender pay gap in EU-SILC (in %), 2013	142
Table A 18: Composition of the explained gender pay gaps at country level in % (2013)	143
Table A 19: Drivers of the occupation-related endowment effect (cross-country estimation) in EU-SILC	146
Table A 20: Drivers of the industry-related endowment effect (cross-country estimation) in EU-SILC	147

Country Codes

Country Name	Code (A2 ISO)
Austria	AT
Belgium	BE
Bulgaria	BG
Croatia	HR
Cyprus	CY
Czech Republic	CZ
Denmark	DK
Estonia	EE
Finland	FI
France	FR
Germany	DE
Greece	GR
Hungary	HU
Iceland	IS
Ireland	IE
Italy	IT
Latvia	LV
Lithuania	LT
Luxembourg	LU
Netherlands	NL
Norway	NO
Poland	PL
Portugal	PT
Romania	RO
Serbia	RS
Slovakia	SK
Slovenia	SL
Spain	ES
Sweden	SE
Switzerland	CH
United Kingdom	UK

Non-technical Summary and Implications

This study undertakes a comprehensive analysis of the sources of wage differences between male and female workers in Europe. Its main purpose is to shed light on the interplay of so far neglected explanatory factors as well as to reveal country differences in the roles of these factors. One specific point of interest concerns the impact of gender differences in the incidence of overeducation. For this reason, our study also examines, in an introductory module, the determinants of overeducation in Europe. In this way, we make contributions to two different, highly debated subfields in labour economics: the overeducation and the gender pay gap literature. In both fields, our innovative features are the large number of determinants as well as the large number of countries simultaneously analysed. The study is divided into three modules, which build on each other. The main findings are summarised in what follows. For more detailed information, please see the full text of the study. For more country-specific information please refer to the **country fiches** provided together with this final report. Some suggestions regarding additional variables potentially increasing the scientific usefulness of the three data sets are made in the last chapter of the study.

As this study is predominantly on the gender divide in earnings, we start with the two modules referring to the measure of the gender pay gap in EU countries. One particular aspect in this context are gender differences in overeducation incidence and returns. This aspect comes last in this summary but first in the main section of the report as it is not restricted to the returns of overeducation but also provides valuable insights into its sources.

Sources of the gender pay gap: results from the SES database

One module (placed second in the report) estimates size and impact factors of the gender pay gap in Europe. First, we provide an update of existing figures on the unadjusted and adjusted gender pay gaps in EU countries based on the most recent wave of the SES (2010). To this end, we performed a statistical decomposition of the observed (unadjusted) gender pay gap. Second, we enrich the literature by undertaking comprehensive country comparisons of the gap components. Overall, we analyse 21 EU countries (plus Norway), which clearly exceeds the scope of existing microdata studies (e.g. Arulampalam et al., 2007; Simón, 2012). Third, we differ from other studies in that we also examine and compare the sources of the part of the gap that cannot be explained by gender differences in worker attributes (henceforth called the adjusted or the unexplained gap), thus providing additional insights into the sources of the pay differential. Finally, we discuss our decomposition results in the broader context of female labour market participation.

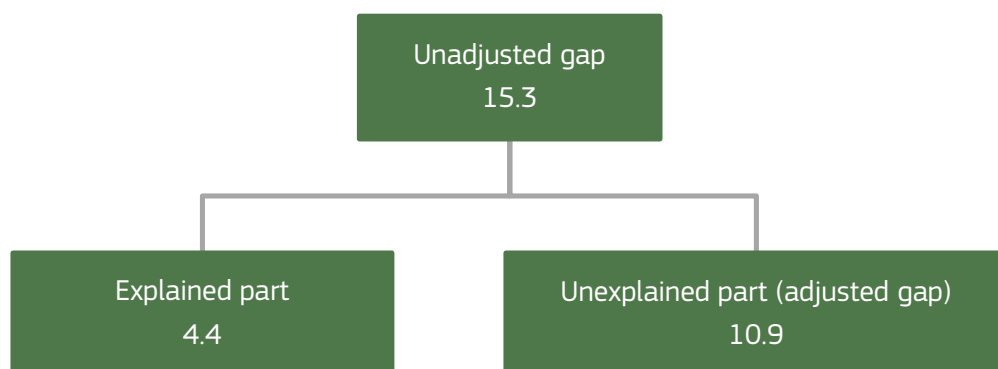
Brief Guide: Decomposition Analysis

The aim of the decomposition technique is to split the gender wage gap into distinct parts. In the Oaxaca-Blinder decomposition applied in this study, the overall (unadjusted) gender pay gap is split into an explained and an unexplained (adjusted) component. The size of the explained component represents the part of the gender pay gap due to the fact that men and women differ in certain observed wage-determining characteristics (e.g. education, occupation). The contributions of the single characteristics are termed the characteristics effects and are determined as the gender difference in the incidence of the observed characteristic weighted with the remuneration of this characteristic for male workers. The size of the unexplained component represents the part of the gender pay gap that is due to the influence of unobserved characteristics and/or due to the fact that men and women receive different returns to the same characteristic. The contributions of single characteristics are termed the coefficient effects.

Decomposition into explained and unexplained gap

As a first result, we measure the cross-country gap in average wages of men and women to be about 15.3% in 2010. This is slightly lower than the most recent result of 16.3% published by Eurostat for the EU-28 in 2013. Apart from the different time of measurement, this discrepancy is explicable by data constraints: in order to account for a sufficient number of explanatory factors in the decomposition, we were – among other restrictions – forced to limit the geographical coverage in our estimations to a subset of 22 European countries (including non-EU country Norway). About one third of the unadjusted gap (4.4%) can be traced back to the role of the explanatory factors included in our analysis. A wage difference of 10.9% remains as the unexplained gap. Hence, women earn 10.9% less than men even when gender differences in educational attainment, occupational choice, working hours and other observable attributes have been taken into account.

Unadjusted, explained and unexplained gender pay gap, EU average (22 countries), based on SES 2010 data, in %



Source: HWWI (2015).

However, at country level the picture varies considerably. The unadjusted gap shows that figures range from 3.6% for Poland to 25.1% for Estonia. From a geographical perspective, it is noticeable that 7 out of 10 Middle and Eastern European states are exhibiting gaps clearly below average, with the Czech Republic, Slovakia and Estonia marking the exceptions. Among the West European countries, only Italy is exhibiting a very small gap (4.5%).

Unadjusted, explained and unexplained gender pay gaps in EU countries based on SES 2010 data, in %

Country	Unadjusted gap	Explained gap	Unexplained gap (adj.)	Country	Unadjusted gap	Explained gap	Unexplained gap (adj.)
Belgium	8.5	2.8	5.8	Latvia	7.4	-3.2	10.6
Bulgaria	8.6	1.1	7.3	Lithuania	5.8	-8.1	13.9
Croatia	5.7	-6.0	11.7	Netherlands	15.2	7.2	8.0
Czech Republic	16.5	3.4	13.1	Norway	14.3	7.6	6.7
Estonia	25.1	10.2	14.9	Poland	3.6	-7.8	11.4
Finland	20.7	9.4	11.4	Portugal	11.4	-0.9	12.3
France	13.5	4.8	8.7	Romania	7.1	0.8	6.2
Germany	22.2	14.5	7.7	Slovakia	16.6	2.2	14.4
Greece	13.1	5.5	7.6	Spain	17.4	5.4	12.0
Hungary	8.4	0.5	7.9	Sweden	14.0	6.3	7.7
Italy	4.5	-6.2	10.7	UK	20.3	6.0	14.3
				<i>Total</i>	<i>15.3</i>	<i>4.4</i>	<i>10.9</i>

Sources: SES (2010), HWWI (2015).

Further country variation is revealed by the decomposition results. The country ranking with respect to the unexplained gap changes substantially compared to the unadjusted gap. Gender differences in average worker features are not only more important in some countries than in others, they also work in opposite directions. For instance, it is striking that the three countries with the smallest unadjusted gap (Poland, Italy, Croatia) all exhibit negative explained gaps. On the other hand, there is a range of countries where gender discrepancies in work-related endowments help to explain significant portions of the overall pay gaps. In Germany and Norway, more than half of the gap can be attributed to this source. In six other countries (Estonia, Finland, France, Greece, Netherlands and Sweden) at least more than one third of the overall gap is driven by this component.

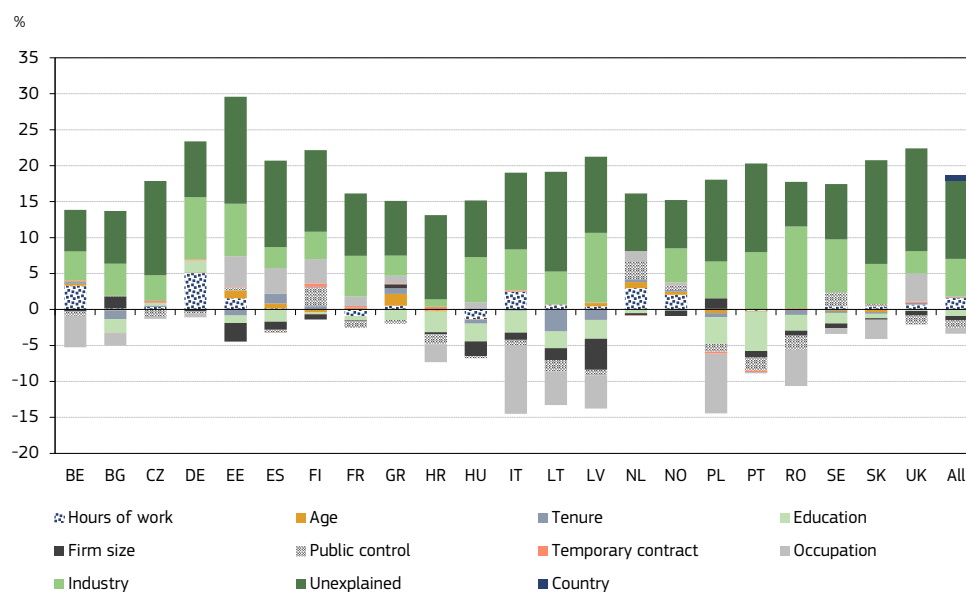
Important context factor at the country level: Female employment selection

An analysis of the relationship of the gender pay gap with female employment rates hints at selection effects as a major source of these country differences. Countries with low wage gaps tend to be characterised by low female labour market participation, suggesting the existence of a positive selection of working women with respect to productivity-relevant characteristics and/or a trade-off between family-compatibility and attractive pay. Therefore, one of our major conclusions is that in discussing wage gaps it has to be kept in mind that a more or less significant part of the female population is not in our sights.

Factors behind the explained gap

The explained gap is further decomposed into the contributions of different worker and job characteristics (see next figure). These are also called characteristics effects: they represent the part of the gap that is due to gender differences in wage-determining features. While some of these features show similar effects across countries, the role of others is highly heterogeneous.

Decomposition of the unadjusted gender pay gap in 2010, in %



Sources: SES (2010), HWWI (2015).

Something that can be noticed for all countries is that the selection of male and female workers into different sectors contributes to the existence of wage differences. Hence, a significant part of the gender gap is due to the fact that women are over-represented in industries with low pay levels. In the classification applied, these are primarily the sectors education, and health and social work activities. Accordingly, men are over-represented in better paid industries. Two of them particularly stand out in terms of a particular high pay level and high male employment share, respectively: construction on the one hand, and chemical products, electric and transport equipment on the other.

Among the remaining characteristics effects, there is none that works in the same direction in each country, while there are some that show a particularly tendency. Three factors can be mentioned that predominantly reduce the pay gap. One of them is firm size. In the clear majority of countries, firms with more than 50 employees exhibit a higher share of female workers than smaller firms. In addition, the payment level in large firms is (under otherwise equal conditions) disproportionally high, implying that this factor mitigates the wage gap. The second mitigating factor is schooling. Female workers in most countries exhibited a higher average level of education than their male counterparts, at least when measured on our three-level scale and this tends to increase wages of females,

thereby contributing to wage convergence. Moreover, the fact that male and female workers are unequally distributed between private and public companies helps to narrow the gap. Publicly owned companies tend to exhibit higher shares of women in their workforce. At the same time, average pay levels in these companies are (all other characteristics being equal) measured to be higher. Finally, there are characteristics that tend to raise the gender pay gap, albeit not in all but in the majority of countries. One of them is hours of work. In all countries under observation, female workers have more often been employed part-time than male workers. In most countries, part-time-work is associated with considerably lower hourly wages. Another channel that tends to widen the gender pay gap is the distribution of temporary vs. permanent contracts. Working in a temporary position reduces the expected earnings in almost all country regressions whereby females are more frequent among temporary workers. Moreover, in the aggregate estimation, the net effect of age differences is practically zero. Effects of the single age groups are of a similar magnitude. A similarly small effect is measured for firm tenure.

Finally, the characteristic causing the most heterogeneous effects is occupation. In Spain and the UK, occupational differences are measured to contribute more than 3.5% to the national overall wage gap, implying this is the prime factor responsible for the existence of a positive explained gap in these countries. In Italy and Poland, we witness a massive negative impact reaching levels of -9.5% and -8.3% respectively, supporting the result that occupational segregation works in favour of women. A detailed analysis of the contributions of the single occupational groups reveals this is mainly the outcome of a strong concentration of female workers in the group of teaching professionals. In both Italy and Poland, the wage bonus of teaching professionals compared to the reference category is much higher than in the cross-country sample. At the same time, the degree to which women are over-represented in this occupational group is higher than in other countries. This observation is again likely to reflect the influence of female employment selection. Being aware of this tremendous heterogeneity on the country level, it does not come as a surprise that the contribution of gender differences in occupational choice to the pay gap at the European level is fairly small. Hence, occupational segregation as such rather takes a back seat and becomes important in interacting with industry and hours worked.

Factors behind the unexplained gap

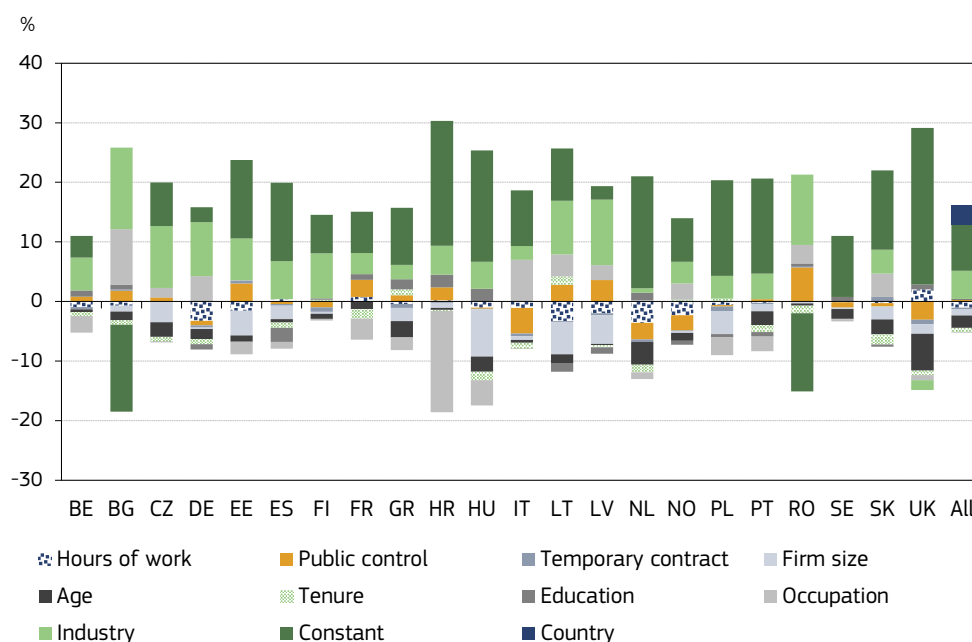
On the country level, the unexplained gap is positive everywhere and makes up the largest part of the overall gender wage gap in almost all countries under observation (with Germany and Norway marking the exceptions). The size of the unexplained component of the pay gap (10.9% for the cross-country sample) represents the part of the gender pay gap that is due to the influence of statistically unobserved characteristics (accounting for 7.7 percentage points) and due to the fact that men and women receive different wage premiums for exhibiting the same observed characteristics (3.2 percentage points). For the unobserved characteristics, gender differences in actual work experience over the lifecycle are expected to make up the bulk of this amount. The wage-reducing effect of a temporary labour market absence of women due to childbirth and childcare is nowhere explicitly accounted for in our approach.

Similar statistically unobserved effects could stem from factors such as personal abilities and different negotiating skills of men and women. For this reason, it is important stress that the *unexplained wage differential must not be equated to the extent of gender discrimination in wage setting*. This does not mean that the sources of the unexplained gap have to remain completely in the dark. As noted

above, the second part of the unexplained gap refers to the different wage returns for women and men to each characteristic (evaluation effect).

Again, some major patterns can be identified. Starting with industry, the gender-specific evaluation of a sector refers to its pay level relative to the reference sector (education). The evaluation effect measures the gender difference in these evaluations. This effect is given particular weight in those sectors with a high female employment share. The fact that in almost all countries (except the Netherlands, Sweden and the UK) the sum of sectoral evaluation effects is positive, contributes to the gender pay gap. For the aggregate sample, this effect equals 4.7%. In Belgium, Sweden and Romania, the magnitude even exceeds 10%. That is, in all countries except the three named above, the pay attractiveness of sectors compared to the reference sector education is stronger for men than for women, hinting at considerable intra-sectoral gender heterogeneity with respect to the sorting into occupations and hierarchical positions within occupations. Apparently, much of the sorting takes place within, rather than in between, industries. An implication is that the wage level effects of different wage setting regimes at industry level (e.g. centralised vs. decentralised, influence of trade unions) in most countries do not fully make up the pay effect of sorting. Moreover, as Goldin (2014) points out, it is likely that sectors with high costs of time flexibility in terms of working hours and temporary employment breaks compensate their employees who stick to the full-time full-year (FTFY) standard with high wage premiums, explaining the within-sector pay inequity with firms' personnel management rationales and not with human capital related differences between men and women in statistically observable *endowments*.

Decomposition of the unexplained part of the gender pay gap (in %), 2010



Sources: SES (2010), HWWI (2015).

The second consistent pattern is the effect of age. Our cross-country wage regressions provide evidence for an inverse U-shaped relationship between age and wage

level for male workers, with conditional pay being highest in the reference group of 40–49 years old. For female workers, instead, conditional wages are estimated to increase continuously over the lifecycle in the cross-country estimation as well as in many single countries. Two reasons for these gender-specific age patterns seem plausible. First, statistically unobserved occupational sorting could again impair our estimations in the sense that women are under-represented in some physically highly demanding jobs in crafts and manufacturing. Second, selection effects associated with the employment decision are another candidate. If with increasing age more and more women come out of the labour market (i.e. become inactive) and the remaining active women represent on average a positive selection in terms of productivity-relevant characteristics, the statistical result of a beneficial age effect emerges.

Sources of the gender pay gap: additional insights from the EU-SILC database

Another module (placed third in the report) aims to reproduce the decomposition analysis of the sources of the gender pay gap in Europe undertaken in module 2 based on an alternative dataset, thereby checking the robustness of our previous results with respect to data coverage and sample selection (see the box *Methodological Guides* for an introduction to this decomposition technique). It also widens the perspective by considering additional explanatory factors for gender wage differences within our decomposition analysis. This provides us with a clearer picture of the forces at work. As part of this model extension, we also examine the effect of overeducation as a potential determinant, supported by results from an additional module investigating the sources of overeducation (see below).

Reproducing the decomposition analysis of gender pay gaps in Europe based on EU-SILC (2013) data has proved to yield valuable additional insights. Making use of this dataset allowed us to add several potential determinants of gender wage differences in comparison to SES. At the same time, we were able to assess the stability of our results from module 2. Looking at the main impact factors, we found some confirmation for our previous argumentation. The sorting of male and female workers into industries with different pay levels is a major contributing factor to the gender pay gap in Europe. Women are over-represented in low-wage sectors such as health and social work activities, but under-represented in high-wage sectors like manufacturing. Gender differences in the number of working hours, especially in the form that women work more often part-time, likewise raise the gender pay gap at the European level.

The only clear deviation is marked by the impact of firm size, now predicting that the average female worker is disadvantaged compared to the average male worker by the fact that she works in a firm of different scale. However, this can be explained by the different composition of the samples. As EU-SILC also contains firms with less than 10 employees, it accounts (unlike SES) for an over-representation of women in very small firms. As small firms on average pay less, women earn less than men since they work more often in small firms.

The effects caused by gender differences in the household-related characteristics introduced with EU-SILC are largely intuitive. The fact that female workers on average declare to be of poorer health magnifies the wage gap, just as the under-representation of women in supervisory positions does. Moreover, we found that the linkages between the pay gap and the phenomenon of overeducation are rather weak, even though we estimated a clearly wage-depressing effect for overeduca-

tion. This is because women in EU-SILC are only slightly more frequently overeducated than men. A worker's place of birth (domestic or abroad) is also of very low explanatory power for the pay gap.

The most interesting information is obtained from the analysis of the partner context. In many countries, being married is associated with higher wages for male workers, which might partly be caused by the existence of marriage-related tax benefits (such as tax splitting) in some EU countries. At the same time, female workers are less often found to be married than males, implying that the distribution of marital status across workers adds to the pay gap. The fact that men and women have partners with different average income and educational attainment has important implications. While the distribution of the educational levels of partners raises the wage gap, the distribution of partner income reduces it. The phenomenon of assortative mating in connection with gender discrepancies in income and education is the likely key for understanding this result. The partner's labour market status is also of relevance. The distribution of inactive partners amplifies gender wage differences, while the distribution of unemployed partners has a mitigating effect. This demonstrates the complexity of channels through which labour market policies can exert influence on the gender pay gap.

Incidence and driving factors of overeducation

Finally, another module (placed first in the report) investigates the incidence of overeducation among workers in the EU and its potential reasons based on data from the European Labour Force Survey (EU-LFS). The concept of overeducation refers to an imbalance between the educational level of a worker and the skill requirements of his or her job. Overeducated workers possess greater educational skills than their jobs require, part of their skills are hence underutilised. If this phenomenon is widely observed, it can have serious implications for the economic prospects of a country. It does not only render public investments into education partly unproductive, it can also harm long-term economic growth by reducing the incentives for future generations to participate in higher education. Shedding light on the roots of this problem is therefore a task of high policy relevance.

Results: Comparison across countries

According to our measurement, the EU-wide share of overeducated workers amounted to 27% for high-skilled and 19% for medium-skilled workers in 2013. These discrepancies are consistent with results of previous work (e.g. European Commission, 2012). At the same time, gender differences in this respect are negligible for medium-skilled workers, while among high-skilled workers men are a little more often found to be overeducated (see graph below). In addition there exists considerable country variation. While about half of the medium-skilled workers are overeducated in Spain and Portugal, in countries like the Netherlands and the Czech Republic the share is lower than 10%.

Share of overeducated workers on the workforce by gender (all countries)

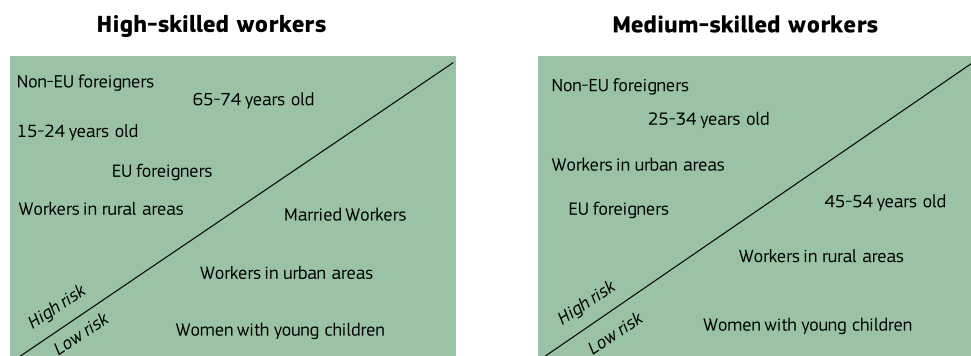
Source: EU-LFS (2013), HWWI (2015).

For high-skilled workers, the picture looks quite different. The overall incidence of overeducation is considerably more widespread in the majority of countries. The only exceptions to this are Spain, Greece, Ireland, Lithuania, Latvia, Portugal and Romania, where the frequency of overeducation is still higher among the medium-skilled. In all, the EU-internal country differences in educational mismatch show that the European community is still far away from reaping the full benefits from cross-border labour migration created by the common market legislation.

Results: Comparison across population groups

In addition to the descriptive analysis, we undertook a regression analysis to quantify the association of a range of characteristics with the probability that a worker is overeducated (for an introduction to the principles of regression analysis, see the box *Methodological Guides* below). First, our analysis shows that the probability of overeducation differs considerably between workers of a distinct age, educational background and gender. High-skilled workers from the youngest (15–24 years) and the oldest (65–74 years) age groups are found to be at a particularly high risk. While the measured high risk for the oldest age group can be attributed to formally low-qualified voluntary activities, the high risk for young tertiary graduates is more a source of concern, especially in relation to the Bologna process. It seems that young graduates have at least temporary difficulties to access adequate positions after graduation. In the role of gender, women per se are not estimated to exhibit a higher or lower risk than men. However, a distinction is made by the impact of family-related decision-making. Married high-skilled women with young children are measured to have a particularly low overeducation probability. In relation, the probability for unmarried high-skilled men is comparatively high. This discrepancy can be understood in the context of the decision to work at all. After starting a family, going to work is associated with higher costs, as less time can be spent with the children. This can raise the minimal wage above which going to work really pays off. Consequently, workers are less willing to work in low-paid, low-qualified positions, reducing their overeducation risk. The fact that this is primarily observed for women indicates that traditional gender roles are still of relevance for employment decisions in Europe.

Risk maps for the overeducation probability of high- and medium-skilled workers



Source: HWWI (2015)

Moreover, we also tested for the effects of other aspects of household composition, namely of the presence of unemployed household members, of inactive members between 15 and 75 years and of elderly members (> 75 years). It has been found that living together with unemployed persons considerably raises the risk of being overeducated. This could be explained by an increased income pressure due to the loss of at least one other source of income in the household. This pressure can force observed workers to accept jobs below their formal level of qualification. Consistent with this, the effects of inactive household members are estimated to be weaker. In most cases, inactivity can be identified as a voluntary out-of-work status, signalling a low-income pressure. Similarly, the effect of elderly household members is negligible. From a policy perspective, these interlinkages point to a deep-rooted connection between the issues of overeducation and unemployment.

With regard to the role of field of study, graduates from tertiary programmes in the fields of teaching and health are measured to exhibit a comparatively low overeducation probability. This can be explained by the risk-reducing effect of educational borders in the access to occupations like teacher and medical doctor. A subject-specific university degree is in both cases a necessary prerequisite for job entry, restricting the field of competitors to those with adequate education. Among graduates from upper (or post-) secondary programmes, the lowest risk is measured for participants of programmes in the fields of engineering, manufacturing and construction. Regarding the effects of job characteristics on overeducation risk, results are largely in line with the existing literature. For both skill groups, risk decreases with increasing job tenure, while it is higher for persons being in temporary employment. For medium-skilled workers, having a second job is also estimated to be risk-enhancing. For high-skilled workers, risk is additionally reduced with increasing hours of work.

Inferences drawn from this study for politics, firms, and social partners

While women approached men in recent decades in terms of formal education resulting in a narrowing of the pay gap, the remaining gap is more a story of within-gender than of across-gender inequality. Female earnings have become more heterogeneous over time, and the same holds for men. Gender differences in working hours and sectoral segregation turn out to mark new earnings divides among genders. Occupational segregation as such rather takes a back seat and becomes

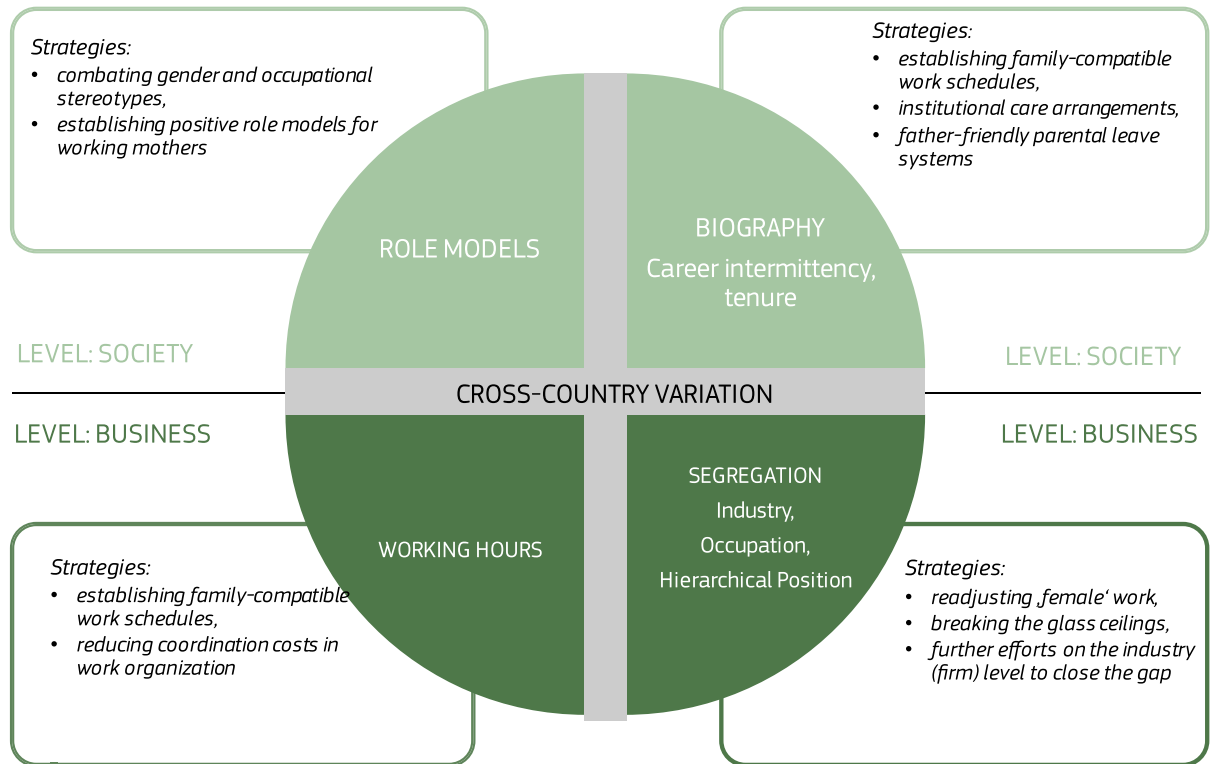
important in interacting with industry and hours worked. For the named earnings divides, firm rationales and technology seem to play a bigger role than gender differences in workers' skills.

Moreover, sectors differ in occupational rewards due to different wage setting regimes at industry level (e.g. centralised vs. decentralised, influence of trade unions). Even more importantly, they differ in their ability and cost to adapt to employees' needs in terms of flexibility, reduced hours and temporary breakouts in the course of family events. Depending on the firm's work organisation scheme, splitting up a full-time job into several flexible part-time jobs or providing flexible hours can create different coordination costs in the cross-firm comparison. As a result, in cost-intensive firms, compensating differentials in hourly payments have emerged, which imply jobs with reduced working schedules are associated with lower hourly wages than classic full-time work (Goldin, 2014). As women and men predominantly differ in their demand for flexible working time arrangements, they are differently affected in their earnings.

In sum, based on the findings of this study and the preceding literature, we can identify four main drivers of the current gender earnings divide. They are listed in the diagram below together with promising strategies for overcoming them, addressing business, politics as well as society as a whole. By shedding light on the role of business, we adopt the recommendation of Goldin (2014), who concluded that promising policies that aim to tackle the gender pay gap should put the sector level more into focus. Decreasing the cost of time flexibility on the firm level is crucial in this regard. Being successful in this sense would enable firms to pay attractive salaries not only to workers who continuously stick to their full-time career over the lifecycle but also to those employees who demand flexible work arrangements in terms of hours and temporary breaks. Lifelong learning opportunities matching changing demands over the life cycle and innovative management tools like shared leadership are essential in this regard. This would be to the advantage of *both* genders. The second aspect on the firm level that deserves more attention is the analysis of prevalent work evaluation schemes with a special focus on a potential devaluation of *female* work. Also, this challenge has to be solved on the firm level and strongly requires the social partners to be involved. Hence, much has to be done in collaborative efforts from different actors.

Other challenges concern the societal perspective. For policymakers, the overarching goal must be to transform the current trade-off between low pay gaps and high female participation rates into a combination of attractive pay and family-compatible work arrangements. Overcoming persisting gender stereotypes as a societal challenge will notably shape the success of this effort. The aim is to free occupational choice from any stereotyped assignment. Establishing positive role models promoting female involvement in leadership positions, gender-untypical occupations and working mothers are promising avenues in this context.

Main drivers of gender earnings divides and corresponding combating strategies



Source: HWWI (2015)

Methodological Guides

Guide 1: Regression Analysis

For a proper understanding of our estimation models and their results, some basic knowledge on statistical regression techniques is a prerequisite. The central aim of a regression analysis is to quantify the relationship between a certain target variable on the one hand and a set of explanatory variables (also called covariates) on the other. This is done by imposing a certain functional form (often a linear function) on this relationship and estimating the parameters associated with this form by means of observed data. Typically, the influence of each explanatory factor is reflected by one such parameter. Signs and magnitudes of the estimated parameter values then allow inferring on the direction and strength of a covariate's effect on the target variable. The special feature of these effects is that they need to be interpreted in an 'all else being equal' manner: they represent the isolated (marginal) change in a target variable y caused by an increase in covariate x by one (marginal) unit, while keeping the entire remaining covariates constant. Hence, regression analysis represents a useful tool for separating and comparing the influence of different kinds of worker characteristics on the wage level or the overeducation risk.

Guide 2: Decomposition Analysis

To interpret the decomposition results for the gender wage gaps in Modules 2 and 3 correctly, some familiarity with decomposition techniques is required. The aim of these techniques is to split an observed discrepancy in some indicator values (e.g. differing wages in euros) between two population groups (e.g. male and female workers) into distinct meaningful parts. In the Oaxaca-Blinder decomposition applied in this study, the overall (unadjusted) gender pay gap refers to men's earnings, that is, it depicts by which percentage women's earnings are lower than men's. The unadjusted pay gap is split into an explained and an unexplained (adjusted) component. The size of the explained component represents the part of the gender pay gap due to the fact that men and women differ in certain wage-determining characteristics (e.g. education, occupation). It is further decomposed into the contributions of the single characteristics, which are termed the characteristics effects. Formally, each characteristics effect is determined as the gender difference in the incidence of the characteristic weighted with the parameter measuring the remuneration of this characteristic for male workers. This parameter is obtained through a preceding regression analysis (see Guide 1). The size of the unexplained component represents the part of the gender pay gap that is due to the influence of unobserved characteristics and/or due to the fact that men and women receive different returns to the same characteristic. Again, a further decomposition into the contributions of single characteristics can be undertaken, which are termed the coefficient effects. Formally, each coefficients effect is measured as the gender difference in wage returns to the characteristic weighted with the average level (prevalence) of the characteristic among female workers.

MODULE 1. Overeducation

1. Introduction

In general, the incidence of 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, both from a macro- and a microeconomic perspective.

From a macroeconomic perspective, an overeducation status of qualified workers reflects a waste of scarce human capital. It has been argued that this misallocation could seriously depress the long-term growth prospects of an economy through the creation of skill-related bottlenecks (Manacorda and Petronogolo, 1998; Sianesi and van Reenen, 2003). Moreover, with respect to public finance considerations, a skill mismatch also implies a direct waste of public means spent on the education of affected workers in the past (European Commission, 2012). Against the background of demographic change, Western economies will start to struggle more and more with the challenge of training a sufficient number of qualified people to meet the growing skill demand driven by trends like digitalisation in the years ahead. This further strengthens the necessity of a productive allocation of scarce human resources. Thus, it will become more and more important to monitor the efficiency of matching processes on labour markets and correct potential misdevelopments through carefully designed policies.

From a microeconomic perspective, the overeducation status can affect a worker's job satisfaction. Overeducated workers are typically found to be less satisfied with their jobs not only compared to adequately matched workers of the same educational level, but also compared to their direct colleagues (Tsang et al., 1991; Battu et al., 2000; Verhaest and Omey, 2006; Korpi and Tåhlin, 2009). 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). These adverse outcomes counteract potential positive spillover effects of knowledge in the workplace (Battu et al., 2000). Moreover, overeducation is associated with earnings losses. Overeducated workers regularly earn more than their job colleagues but less than correctly matched workers with similar education (e.g. Daly et al., 2000; Bauer, 2002; Boll and Leppin, 2014a).

In all, the current stand of research suggests that overeducation entails considerable risks for the prosperity of an economy both in micro and macro perspectives. However, before being able to tackle the problem successfully, it is essential to understand the driving forces of overeducation at the individual level. Thus, the

aim of this paper is to identify possible determinants of overeducation for EU-28 countries. We make use of wave 2013 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. Furthermore, we carry out separate estimations for high- and medium-skilled workers to account for potential differences in the channels leading to overeducation.

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, focusing on household characteristics: the presence of inactive persons in the same household, the presence of unemployed as well as that of persons older than 75 (as a proxy for eldercare). Moreover, the effect of having a second job is added as another potential job-related factor. Second, our results allow for the first time 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. Finally, by undertaking estimations differentiated by skill level, we are able to analyse how far certain channels affect the overeducation risk of workers at different educational levels differently. This is of particular relevance for factors like field of study, whose interpretation is strongly related to educational status.

Our findings indeed point to a considerable variation in the potential determinants across countries as well as across skill levels. This variation is not restricted to job-related characteristics, but interestingly also concerns household variables. Among those determinants showing a largely uniform influence are nationality, job tenure, temporary employment and presence of unemployed household members.

The rest of the paper is organised as follows. The next section provides a brief overview of the theories and empirical results regarding the determinants of overeducation. Section 3 describes the measurement method and the data, and section 4 the model set-up. The results are discussed in section 5 and section 6 concludes.

2. Literature Findings

2.1 Theories

In principle, a vertical skill mismatch can result from characteristics of the worker, characteristics of the job or characteristics of the worker-job matching procedure. The first case is emphasised by supply-side theories of labour productivity. According to the neoclassical *human capital theory (HCT)*, a worker's productivity on the job is first and foremost determined by her past investments into human capital through formal education or training. These investments are under normal circumstances rewarded by the market, as workers get paid according to their marginal product. The HCT therefore regards job mismatch as a negligible and temporary phenomenon, which is naturally corrected by the market (Becker, 1964). Any short-run oversupply of skilled labour would be eliminated through wage adjustments and a corresponding change of factor intensities in production. Based on this theory, the widespread observation of overeducation in job matching could be interpreted as an indicator either for a disequilibrium situation or for the existence of heterogeneous abilities across workers. If job-relevant abilities do not perfectly correlate with education levels, the latter can no longer be viewed as appropriate measures for qualification levels. A lack of formal education could to some degree be compensated by a high level of talent or work experience. This can also hold vice versa: a lack of job-related ability requires some compensation in terms of additional human capital as signalled by an educational degree (Korpi and Tåhlin, 2009). In this respect, overeducation does not have to imply overqualification, but rather signals a lack of those skills unrelated to schooling. The implication is that observed overqualification is simply a measurement error due to the presence of statistically unobserved differences in abilities or worker preferences (e.g. preference for family-friendly work schedules). Chevalier (2003) names this measurement error 'apparent overeducation', to be distinguished from 'genuine overeducation'. In any case, in terms of productivity, the theory postulates that in fact, all workers are correctly matched.

This kind of reasoning rests on the assumption that skills gathered through schooling and skills gathered through practical experience (or inherited talents) are highly substitutable in shaping a worker's productivity. This clearly does not fit reality for those occupations where a certain level of experience is essential for promotion. Against this backdrop, the *career mobility theory* was developed as another attempt at a supply-side explanation (Sicherman and Galor, 1990; Sicherman, 1991). It views the entering of workers into a skill mismatch as a deliberate decision made to improve their long-term earnings prospects. By gathering experience at low entry levels, labour market entrants can enhance their skills and raise their chances of occupational upgrading. Under the precondition that graduates with higher education levels face a higher probability of promotion, this strategy can serve as an explanation for the occurrence of overeducation. Individuals may choose an entry level in which the direct returns to schooling are lower than those in other feasible entry levels if the effect of schooling on the probability of promotion is particularly high in the given firm. Similar to HCT, the implication is that overeducation should not be seen as a cause for concern, but merely as a temporary phenomenon con-

sistent with efficiency goals. From the perspective of an individual worker, it represents a reasonable investment to improve future employment opportunities. One underlying assumption is, as in the case of HCT, that the job market can adapt in a flexible manner to changes in the provision of skilled labour.

In contrast, the *theory of job competition* first formulated by Thurow (1975) assumes that such an adaptation is excluded, at least in the short-run, due to a rigidity of the wages paid for specific occupations. Workers therefore compete for jobs in certain occupations, not for wages. The set of jobs, in turn, is the result of firms' production requirements. The higher an individual's formal level of education, the lower is its expected cost of training. As a consequence, more educated people will have a higher chance of securing a job in a certain occupation. Unsuccessful applicants will switch to less demanding jobs. If there is a shortage of jobs for professionals, some high-skilled workers will have to be satisfied with low-skill jobs. This is a demand-side explanation for skill mismatches. Apart from the level of education, work experience also matters in determining an individual's place in the job queue. The group of overeducated workers will therefore primarily consist of the less experienced (or generally less able) individuals among the skilled workforce.

The *assignment theory*, in turn, seeks to unify both supply- and demand-side explanations (Sattinger, 1993). Based on the *theory of job search* (Jovanovic, 1979), it views overeducation as an inefficient outcome of a job matching procedure. Due to the existence of search costs, highly educated jobseekers might be satisfied with finding a position at a level below their formal qualification. At the same time, employers are happy to hire applicants whose skills exceed the current job requirements, as this could allow them to save on training costs in the future. In addition, a skill mismatch in this framework can also arise through heterogeneous preferences among workers. Wages and qualification levels are not the only indicators in the search process. Applicants typically also have preferences and dislikes for certain occupations. These influences can induce them to choose matches in which they are paid below their productivity. On the other market side, employers might also have specific preferences for some applicants beyond formal qualification levels. As the true skill level is not fully observed, it can be perfectly rational for firms to screen workers for other productivity signals such as the field of study and non-education related activities. Matches characterised by overeducation are thus explained by incentives both on the supply and demand sides of the labour market. In opposition to human capital theory, overeducation is thus considered a permanent phenomenon. Moreover, it is also considered to be more worrying than the previous theories, as any degree of excess education (i.e. exceeding the requirements of the respective job) is predicted to yield no additional returns to the worker, implying that some portion of human capital investments is turning to losses.

In this context, a special application of a job-search framework is the *theory of differential overqualification* developed by Frank (1978). He attempts to explain the frequent occurrence of overeducation among women, particularly the married ones, based on matching problems. His starting point is the notion that married couples cannot optimise their job choices independently of each other, as they face a co-location restriction. In traditional gender role model settings where the couple's priority is on the job match of the husband, the husband acts as a first-mover, i.e. he performs his job search first. Given that he faces no spatial restrictions in the first place, he could principally search within a large market area and thereby increase the chances of finding a good match. After he has accepted a match, the wife will conduct her job search. However, due to the co-location restriction, she can do that only within a much smaller market area. The likelihood of finding a job adequate to

her qualification level is therefore much lower for her than for her husband, explaining a striking incidence of overeducation among married females. This deprivation of married women is predicted to be the stronger, the higher their mobility costs. Büchel and van Ham (2002) extend this theory to all job seekers: those people located in regions with small regional labour markets and high unemployment face the largest risk of overeducation.

The theories mentioned so far are definitely the most popular ones. A further commonality is their focus on individual optimisation as a driving force and a corresponding neglect of the important role of institutions in job matching. Relevant institutions are the educational system on the one hand and the labour market structure on the other hand (Shavit and Müller, 1998). So far, the role of the former has been analysed more intensively than the role of the latter. Specifically, the degree of stratification (diversity of tracks within the educational system) has received attention as an explanatory factor of match quality. Highly stratified systems of education with strong apprenticeship systems usually generate a lower number of upper secondary and tertiary graduates (Shavit and Müller, 1998). Highly stratified systems are thus expected to suffer less from overeducation, as they reduce the intensity of competition among highly educated workers for top jobs. Furthermore, the level of standardisation of educational degrees within a country is predicted to be negatively associated with overeducation risk. A strong homogeneity of degrees implies low barriers to labour markets outside the own region and thus raises chances of finding an appropriate match.

2.2 Empirical results

With regard to the empirical testing of these alternative theories, the career mobility hypothesis has, after its initial postulation, received the least amount of confirmation. For instance, Battu et al. (2000), Büchel and Mertens (2000) as well as Baert et al. (2012) were unable to detect significant future wage growth for overeducated workers. This rejects the view of overeducation as the starting point for a promising career. Baert et al. (2012) even find evidence for the existence of an overeducation trap where workers systematically get stuck in low-paid jobs underutilising their skills. Insufficient access of overeducated workers to training activities is identified as one channel through which such a trap can emerge.

On the other hand, the human capital theory and the job competition theory both receive some confirmation through the negative impact of work experience on overeducation risk estimated by most studies (e.g. Alba-Ramirez, 1993; Büchel and Battu, 2003; Nielsen, 2011, Boll et al., 2014b). The human capital theory, however, is incompatible with the simultaneous observation of overeducation and an increase in the labour market returns to skill. To be consistent with the view of overeducation as a short-term disequilibrium phenomenon, an opposite evolution should have been expected. The job competition theory, in turn, is clearly at odds with the result of a positive return to surplus education for overeducated workers (e.g. Duncan and Hoffman, 1981): labour productivity is thus not purely job related. Among all general frameworks, the assignment theory seems currently to be the empirically most well founded. However, this is also the concept that allows for the fewest general conclusions. Finally, the theory of differential overqualification is facing mixed empirical evidence. While McGoldrick and Robst (1996) reject the hypothesis that married women face a significantly higher probability to be overeducated, Büchel and Battu (2003) find some evidence for this fact after controlling for commuting distances.

Individual characteristics

To some extent, the risk to be overqualified can be attributed to an individual's characteristics. This concerns both innate characteristics and those that are results of the individual's education and employment history. Among the innate characteristics, the role of gender differences has received a large amount of attention in recent literature. This interest primarily arises from the observation made in many countries that the share of overeducated workers among women is higher than among men. One question is whether this observation is primarily the result of gender discrimination or of gender-specific differences regarding industry choice, educational background or work experience. Econometric studies therefore attempt to isolate the gender effect, usually by controlling for a large set of related variables. Despite the considerable amount of empirical work, no clear picture on direction and significance of this effect has yet emerged. A slight majority of the studies examined finds that the effect of gender on overeducation risk is insignificant in a multivariate set-up (Büchel and Pollmann-Schult, 2001; Groot and van den Brink, 2003; Frenette, 2004; Green and McIntosh, 2007; Capsada-Munsech, 2015). Alba-Ramirez (1993), Groot (1996) and the European Commission (2012) obtained the result that male employees face a slightly higher overeducation risk, an effect which is however in all cases only weakly significant. In contrast, Ortiz and Kucel (2008) estimate that female workers are at significantly higher risk. These contrasting results could at least in some part be explained by differences in the measurement methods applied. McGoldrick and Robst (1996) point at this possibility by comparing estimation outcomes based on an objective measure of overeducation (statistical distribution) with a subjective one (questionnaires). They find that women are significantly more likely to experience overeducation than men under the subjective measure, but significantly less likely under the statistical measure. For this, they offer the explanation that women on average tend to be assigned the less attractive jobs within a statistically defined occupational group. This gendered pattern of overeducation incidence is confirmed by a study of Boll et al. (2014b) for West and East Germans in the time span 1992–2011, though for West Germany only with respect to self-assessed overeducation. In terms of statistical overeducation, West German female graduates recently outperformed West German male graduates and West German women with medium education mostly exhibited higher rates of overeducation than their male counterparts in the period under observation.

Moreover, one strand of literature focuses on the interaction of gender with marital status, mostly based on the theory of differential overqualification outlined above (for the interaction with the number of children, see household characteristics). The tested hypothesis is that within the group of married workers, women face a higher overeducation risk as their job search is limited to a geographically smaller market area. The available evidence for this idea is rather weak. Frenette (2004) finds in his sample of university graduates no significant differences in overeducation risk for married females either two or five years after graduation, in comparison to unmarried females or married males. Groot (1996) likewise detects no significant impact of marital status for women, but interestingly a risk-enhancing effect of marriage among men. In Sloane et al. (1999), interaction in terms of gender and marital status are also found to be insignificant. However, the authors argue that some indirect evidence for differential overqualification could be seen in the positive and significant effects of the variables spouse in management and spouse in self-employment, obtained when estimating their model separately for women. When interpreting these as indicators for labour market success, it could imply that a spouse's successful career limits the wife's opportunities for her own advancement, which is in line with Frank (1978).

In contrast, McGoldrick and Robst (1996) conduct an alternative indirect test by introducing a term interacting gender with the size of the local labour market (proxied by population size) into their model. The insignificance of this term for a subsample of married workers indicates that a smaller local labour market does not raise overeducation risk for married women. Büchel and Battu (2003) pointed out that this approach did not account for gender differences in maximum commuting distances. In their sample from the German Socioeconomic Panel, they find that married women in rural areas face a significantly higher risk of overeducation than single city men. However, when introducing commuting distance into their model, being located in a rural area starts to raise the overeducation probability for both married men and married women, which is at odds with the argumentation of Frank. This result was confirmed by Ramos and Sanromá (2011) based on Spanish data. Finally, Joonas et al. (2012) in their estimation using Swedish register data even obtained the result that the fact of being married reduces overeducation risk for the subsample of women. However, this seemingly contrary evidence has to be interpreted cautiously, as it does not account for the impact on a woman's decision whether to work at all. Women facing a lack of sufficiently adequate matches might prefer to turn inactive rather than accepting a job below their skill level. Part of the measured lower risk can thus be viewed as the outcome of a selection process concerning labour market participation. Büchel and van Ham (2002) document the impact of this selection process for German workers.

Another potentially relevant individual characteristic is the worker's age. Here, the question is to what extent age can be seen as a genuine determinant and to what extent it merely serves to approximate experience-related skills. As work experience commonly shows a strong positive correlation with age, it could be expected that older workers are less frequently overeducated due to the skill-enhancing effect of increased experience (see work experience below). However, there are also other reasons for age-based discrimination. For instance, independent of actual work experience, older workers tend to be more accustomed to the process of job searching. They might have a better overview on local market conditions than first-time entrants and thus also superior information on the relative quality of a certain match. As a consequence, they might be less prone to the risk of accidentally being trapped into a low-qualified position despite better opportunities elsewhere. On the other hand, with increasing time spent on the labour market, the risk increases that certain skills gathered through formal education have become obsolete in the meantime. Strictly speaking, this does not need to affect the probability of overqualification, as it is accompanied by a depreciation of human capital over time. Nevertheless, it raises the probability of overeducation, as this measure typically applies to the formal education level.

In comparing the empirical results on the age effect, it is for the above-mentioned reasons essential to distinguish whether an approach includes work experience as a separate determinant or not. Given that the high degree of correlation between the two measures implies a risk of biasing the results, it is a general practice in the literature to either consider one or the other. Within the first group, the European Commission (2012) finds a continuously shrinking overeducation probability for categories of workers of increasing age. Büchel and van Ham (2002) estimate a significant effect only after controlling for the selection effect of employment decisions. In this specification, the effect is generally positive, implying a higher overeducation risk for older workers. This result might partly be explained by the decision of old workers that suffer from a skill decline to leave the labour market prematurely. Their increased overeducation risk is only captured when controlling for sample selection. Ortiz and Kucel (2008) likewise estimate a significantly positive effect both for Spain and

Germany, but again not for all specifications. In contrast, Groot and van den Brink (2003) detect no significant impact of age on the incidence of overeducation in any of their model types. Another open question concerns the linearity of the age effect. In wage regressions, age often enters the equation as a quadratic term, which is supposed to reflect the inversely U-shaped relationship of productivity and years of living. As a consequence, a U-shaped pattern for the change in overeducation risk with increasing age could be expected: while up to some point workers benefit from increased experience, the risk that certain schooling-based skills become outdated continuously gains importance. The available scarce evidence provides at least some confirmation for this hypothesis. Green and McIntosh (2007) as well as Joona et al. (2012) integrate an additional quadratic term for age into their models. Both obtained a significantly negative coefficient for the linear and a significantly positive coefficient for the quadratic term. The effect thus changes its sign in the expected direction. In contrast, both linear and quadratic terms are insignificant in Büchel and Battu (2003).

Those papers focusing on the impact of work experience establish a more clear-cut picture. By far the most dominating outcome is a highly significant negative impact of increased experience on the incidence of overeducation (Alba-Ramirez, 1993; Groot, 1996; Nielsen, 2011; Sloane et al., 1999). In Büchel and Pollmann-Schult (2001), the same effect is only weakly significant. In Boll et al. (2014b), the significance level varies between subsamples differing in region and educational attainment as well as between the chosen measures of overeducation. The most striking exception is McGoldrick and Robst (1996), where the effect is insignificant for all three measures of required schooling.

A seldom analysed determinant is field of study, in some part 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 offer more occupation-specific skills, raising the chances for graduates to find appropriate jobs in the corresponding occupational groups. 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. Second, 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 mathematics or natural sciences, which enjoy the reputation of imposing high intellectual demands on their students, could convince employers of the extraordinary talent and/or motivation of applicants. This could give them preferred access to positions with high skill requirements, possibly also outside the occupational groups associated with their subjects.

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 programmes. 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 subsample of university graduates and thus the tertiary level. They make a quite detailed distinction among 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 maths is explained by 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, business 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 higher than for the reference were also 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 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 comparability is limited by the different field classifications, the literature results suggest some considerable degree of heterogeneity, with students of social sciences and humanities being at higher risk than those in natural and related sciences.

Household characteristics

In addition to a worker's individual background, part of her overeducation risk can also be related to her living situation. The literature so far has focused on the presence of children as a determinant. Childcare requires a perpetual allocation of resources in the form of time and/or money. In this way, taking a job creates additional opportunity costs that can influence a parent's decision on which job match to accept. In general, both number and age composition of children could be considered relevant indicators of these opportunity costs. The higher the number of children and the lower the age of the youngest child, the higher are the expected opportunity costs. The specific linkage of these factors to the likelihood of overeducation is nevertheless complex. One channel of influence might be the impact on a worker's reservation wage. A high reservation wage due to high opportunity costs can induce a job applicant to turn down low-pay offers with low qualification requirements, thereby reducing the frequency of overeducation. On the other hand, some comparatively less demanding jobs especially in administration allow for more time flexibility than most high-level leadership positions. This fact raises the attractiveness of these jobs for parents, triggering an opposite effect of children on overeducation risk. In the same vein, the need to take care of children can limit the spatial mobility

of parents and thus the number of appropriate job matches available within their restricted area of job search.

In line with these thoughts, the existing evidence shows no uniform picture. The literature so far has set its focus on number of children as a household indicator, commonly including an interaction term with gender. Büchel and Pollman-Schult (2001) limit their measurement to pre-school children (<7 years) and consider an interaction of children number with female sex. This does not deliver significant coefficients for any specification. In contrast, Büchel and van Ham (2002) include all children up to 16 years and also control for the selection effect of the employment decision. Without controlling for self-selection, they find for female workers, a significant positive impact of the number of children on the risk of overeducation. Significance however disappears in the Heckman self-selection specification, indicating that the most relevant effect of children already influences the decision to enter the labour market. Green and McIntosh (2007) merely test for the impact of having children or not. It proves insignificant as long as job-related determinants are included in the model. When they are omitted, the children dummy significantly raises the overeducation probability, which signals a strong linkage between family situation and job type.

Several family-related aspects are as yet insufficiently explored. First, the role of a child's age deserves more attention. Sloane et al. (1999) provide at least some insights in this regard. They distinguish between children up to 2 years and children between 3 and 5 years in their model specification. While the number of children younger than 2 years exerts a significant and positive impact on the overeducation risk of women, the effect for older children remains insignificant. This is in line with the idea that younger children impose tougher restrictions on the employment decisions of their parents. Furthermore, the impact of the presence of other household members is also worth investigating. For instance, living together with one or more elderly dependents could be expected to influence job search activity and thus also the overeducation probability. For the direction of the effect, the same set of opposing arguments could be mentioned as for the presence of children. To the best of our knowledge, this channel has so far not been accounted for in the overeducation literature. Moreover, the presence of non-working household members within the remaining age category of 15–74 years could also induce a relevant distortion of labour supply. This includes on the one hand active job seekers coded as unemployed, on the other hand inactive persons outside the labour force (e.g. students, housewives and husbands, chronically sick people, retirees younger than 75). Given that these two groups are also likely to differ in their contributions to the household income, it seems worthwhile to estimate their impacts on the overeducation probability separately. The sign of these impacts is a priori unclear. However, it has to be noted that by their search activities and availability for work, the unemployed prove to be more attached to the labour market than the inactives. Hence, the latter are more likely to be deliberately jobless than the former. As a consequence, the additional earnings burden levied on the employed household member should be lower in the case of an inactive than in case of an unemployed household member for two reasons. First, the household's earnings might be regarded as sufficiently high, second, the inactive person might be able to adapt short-term to a changing income situation by switching to an additional earner status. The need to contribute to the subsistence of others might give rise to a preference for high-wage occupations with correspondingly high skill requirements. On the contrary, one could also argue that the same need might force applicants to accept rather unsuitable positions instead of waiting for better matches. Moreover, a worker's statistically unobserved

productivity might also be linked to the composition of his or her household. Likewise, this is an area that remains to be explored.

Job-related characteristics

Previous studies demonstrate that the incidence of overeducation is strongly related both to job type and firm characteristics. With job type, one relevant distinction relates to contract length. Economic logic would suggest that people with fixed-term contracts are more likely to work in positions for which they are overeducated than people with permanent ones. Due to the transitory nature of fixed-term jobs, workers are less concerned about qualification levels, as they tend to view these matches as merely temporary solutions on their way to more favourable permanent positions. Green and McIntosh (2007) as well as Ortiz (2010) indeed find some evidence for a significantly lower overeducation risk among workers in permanent positions.

Less obvious from a theoretical point of view is the relation to the extent of work. In this regard, one channel of influence is the wage rate. In a world where people were allowed to freely choose their workload, they would respond to receiving a bad job match with low skill requirements and a low hourly wage by adjusting the amount of hours worked. Whether they would decide to work more or less than in better paid jobs is however ambiguous, due to the presence of contrary substitution and income effects. Given that flexibility is in reality limited, a more frequent situation is where applicants have to choose between jobs associated with standardised workloads, i.e. full-time or part-time. In this situation, several arguments would suggest a lower overeducation risk for full-time workers. First, for people choosing to work part-time, their job-related activities make up a considerably smaller share of their entire daily activities than for full-time workers. Therefore, their willingness to execute job tasks that do not fully match their qualification levels can be expected to be greater. Moreover, jobs with longer working hours can also create better opportunities for training participation and advancement, thereby improving the match quality over time. An inverse channel could arise from the impact of the overeducation status on the willingness to work overtime: the literature has established a negative association of overeducation with job satisfaction, which could reduce personal engagement in this regard. All these arguments point to a negative correlation between hours of work and the overeducation probability. Green and McIntosh (2007) indeed document such a relationship for workers in Great Britain through a positively significant part-time dummy. Moreover, Groot and van den Brink (2003) in the Netherlands come to the same result when measuring work extent continuously (in hours of work). In contrast, the European Commission (2012) does not detect a significant association with workload.

The empirically best documented job-related determinant of overeducation risk is job tenure. It measures the length of the current employment spell, commonly identified as the amount of days spent with the current occupation in the current firm. It is a concept closely related to questions of job mobility and in this way also to the quality of job matches. A long tenure signals a high level of satisfaction with the match received. As job satisfaction is in turn linked to the incidence of overeducation (e.g. Allen and van der Velden, 2001), one would expect a long tenure to imply a low likelihood of overeducation. Another channel can emerge through internal mobility of workers within firms, as suggested by the career mobility theory. The longer a worker stays in a firm, the higher is the likelihood of advancement into better positions with higher skill requirements and thus lower overeducation risk. This

is confirmed by the existing evidence. A wide range of papers detects a significant negative effect of job tenure on overeducation risk in a variety of datasets (Sloane et al., 1999; Büchel and Pollmann-Schult, 2001; Büchel and van Ham, 2002; Büchel and Battu, 2003; Groot and van den Brink, 2003; Ortiz, 2010; European Commission, 2012). An exception is Groot (1996), who measures the association to be significantly positive. He interprets this as an indication for a negative relationship between tenure and individual productivity. Low-productive workers receive fewer job offers and therefore tend to remain stuck in bad matches underutilising their skills. Whether relevant or not, this phenomenon seems to be outweighed by opposing effects in other studies.

Finally, some evidence exists on the impact of training activities. The main purpose of on-the-job training is to enhance a worker's occupation-specific qualification. Therefore, one could expect that more intensive training participation in the past reduces the likelihood of overeducation today. Moreover, employers might be less willing to invest in additional training of workers which are already formally overeducated, likewise suggesting a negative coefficient for training. On the other hand, voluntary training participation could also be a sign of dissatisfaction with the current job situation, potentially resulting from an existing skill mismatch. This would point to an opposite relationship. The empirical literature has so far unanimously found a negative association with overeducation risk (McGoldrick and Robst, 1996; Büchel and Pollmann-Schult, 2001; Büchel and Battu, 2003; Capsada-Munsech, 2015).

3. Measurement and Data

An empirical identification of overeducation is aggravated by two basic problems. The first one is its strong dependence on the judgement about the skill requirements of a particular job. As the job type is often classified based on occupational groups (which consist of a range of more specific occupations), some degree of unobservable heterogeneity remains. This threatens the validity of the skill assessment and thus potentially biases the measurement of overeducation. Second, another form of heterogeneity relates to statistically unobserved personal abilities, which influence the true skill level of a job applicant (Battu et al., 2000). This implies that overeducation and overqualification should be regarded as two distinct phenomena. Overeducation might simply be a sign for a weak correlation of educational degree and personal talent (see above). This blurs the interpretation of the measure as an indicator of skill mismatch.

The first step in measuring the extent of over- or undereducation is to assign an appropriate educational level to a certain job. The approaches applied by the literature can be divided into three different methods. The first one is an objective method based on a systematic evaluation of the educational level required for an occupation, made by a group of experts. This rests on the implicit assumption that job characteristics within an occupation do not vary. Without a continuous updating of the assignments, it can also create a bias by not considering technology-related changes in the skill requirements associated with certain occupations. For instance, the negligence of growing skill demands could contribute to an overestimation of overeducation in the workforce.

The second method is a subjective approach that relies on workers' self-assessments of the educational levels required for their specific jobs, typically attained through questionnaires. The obvious advantage of this method is that it is independent of a formal definition of occupational groups, therefore not subject to the measurement risks mentioned above. One criticism is that workers might have insufficient information about job characteristics and the labour market composition.

Finally, a third method derives a statistical measure of overeducation based on the data observed (realised matches approach). By taking averages of the actual levels of education observed for workers in an occupational group, overeducated workers can be identified as positive outliers in the distribution. To determine a threshold defining outliers, a standard deviation measure is prevalent. It identifies normal education levels for an occupation as those lying with one standard deviation distance around the mean (Verdugo and Verdugo, 1989). One drawback of a mean-centred approach is its high sensitivity towards outliers. Kiker et al. (1997) propose to amend this by choosing the modal value of the distribution instead of the mean as the reference point. Compared to expert assessments, the advantage of statistical approaches is their responsiveness to technological change. On the other hand, no systematic overeducation in certain branches can be analysed, as assessments are merely based on the relative position within the distribution. Moreover, both versions of the approach can yield unreliable estimates when sample sizes for certain occupations are low (Battu et al., 2000).

We used data from the European Labour Force Survey (EU-LFS)¹ to identify possible determinants of overeducation. The EU-LFS is a quarterly household sample survey, including the 28 Member States of the European Union, two candidate countries (Macedonia and the former Yugoslavia Republic) and three countries of the European Free Trade Association (Iceland, Norway and Switzerland). It has been based on European legislation since 1973 in order to assure comparability of the results. The survey design, survey characteristics, methods and the decision-making processes of the EU-LFS are precisely regulated. The EU-LFS covers approximately 1.8M individuals aged 15 years or older² and asks the respondents for their demographic background, labour status, employment characteristics and their previous employment experience/search for person not in employment. Furthermore, respondents are asked whether they were students or participated in any courses during the past four weeks, their field of study and their highest educational attainment level. Respondents are asked via personal visits, telephone interviews, web interviews and self-administered questionnaires. In 2012, the participation was compulsory in thirteen countries (Belgium, Germany, Greece, Spain, France, Italy, Cyprus, Malta, Austria, Portugal, Slovakia, Norway and Turkey), and voluntary in the other countries. The quarterly sampling rates in 2013 varied from 0.25% in Germany and France up to 4.5% in Luxembourg. The overall sampling rate of the total population was about 0.3%.

Our analysis is based on 2013 data and is restricted to the EU-28 countries. Respondents are assigned to countries based on their place of work. Malta, Poland and Slovenia are excluded from the analysis due to the lack of detailed information regarding occupation groups. Furthermore, the sample is restricted to respondents aged 15 to 74 years. Although we focus on highly educated workers/graduates, we also estimate our models for medium-skilled workers and compare the results. The highly educated have completed tertiary education (5A, 6), and the medium educated people have completed upper secondary or post-secondary education (ISCED 3-4, 5B). We use the EU-LFS weighting variable (COEFF) in order to calculate the modal qualification level of each occupation group.

We refer to the above-mentioned overeducation as a vertical inadequacy. Therefore, we use the realised matches approach 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.³ To investigate the causal factors of overeducation, we use a broad range of explanatory variables. In particular, we differentiate between three categories of covariates, namely personal characteristics, household characteristics⁴ and job characteristics. Moreover, we consider interaction terms between sex and different household characteristics. Detailed summary statistics on all variables are provided in Tables A1 and A2 in the Appendix.

Personal characteristics include gender, 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 use the following six age dummies:

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

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

3 The use of one of the alternative methods mentioned above is impossible with the given dataset.

4 Household characteristics are not available for the Nordic countries (see European Union, 2014).

15–24 years, 35–44 years, 45–54 years, 55–64 years and 65–74 years. Those aged between 25 and 34 years belong to the reference group. Note that a person's age is only given in age groups comprising five years each (e.g. group 0–4 years) in the anonymised version of the EU-LFS. Since this study aims to identify the interplay between overeducation and field of highest level of educational attainment, our models also include a person's highest level of educational attainment. Therefore, we differentiate between eight broad fields of study: teacher, training and education science; humanities, languages and arts; social sciences, business and law (reference group); science, mathematics and computing; engineering, manufacturing and construction; agriculture and veterinary; health and welfare; services. Respondents who classify themselves into general programmes are excluded from the analysis, as these programmes usually do not lead to a generally acknowledged degree.

Referring to household characteristics, we control for the number of unemployed and inactive adults, the number of people aged 75 and over (eldercare) and the number of children (between 0 and 5 years, between 6 and 11 years and between 12 and 14 years) in the same household. Due to data restrictions these variables are not included in the models for the Scandinavian countries (Denmark, Finland and Sweden).

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 his or her main job. Tenure is defined as the number of years since a person started to work for his/her current employer or as self-employed. Both variables are also included as quadratic terms in our models. Further job characteristics are considered by means of dummies that are equal to one if the respondent is usually working less than 15 hours per week (indicator for marginal employment), if he or she holds a temporary contract or if he or she has a second job, respectively. Firm size is controlled for by means of 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. Whether a respondent attended any courses, seminars, conferences or received private lessons or instructions outside the regular education system within the last 4 weeks prior to the interview is captured by a lifelong learning dummy. The purpose of these most recent taught learning activities can be either job related or mostly personal.

As a variable reflecting the spatial dimension, the degree of urbanisation is included. It varies from 1: densely populated area to 3: thinly populated area and is available for all countries. We also include economic sector and country dummies in our regressions. By including country dummies we control for different country sizes. Hence, the cross-section results may be interpreted as average effects of country-specific regression results.

Finally, we include interaction terms between gender and age, field of study as well as a number of personal and household characteristics: marital status, the number of children, the number of people aged 75 years and older and the degree of urbanisation. We do this to account for likely discrepancies in the marginal effects of these variables between male and female workers, as suggested by theoretical reasoning (see section 2).

4. Model Set-up

In order to estimate the probability of being overeducated we make use of a probit model as a specific form of a regression model (see Methodological Guides for an introduction to regression models). The following section briefly describes this type of model. For more details, see Judge et al. (1988). A probit model is a binary response model where a respondent can either be overeducated or not, in this case.

Assume an unobservable/latent variable y_i^* which can take all values in the range $[-\infty, +\infty]$. The underlying latent model is given by:

$$y_i^* = X_i\beta + \varepsilon_i.$$

Unfortunately, we do not observe y_i^* , we just observe whether a respondent is overeducated or not:

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ 1 & \text{if } y_i^* > 0 \end{cases}$$

Hence, the probability that someone is overeducated is a function of the independent variables and is given by:

$$p = \Pr(y_i = 1 | x) = F(x_i\beta).$$

Since y_i^* is unobservable, we do not know the distribution of the errors ε_i . Regarding the probit model, ε_i is assumed to follow a standard normal distribution of the following form:

$$F(X\beta) = \Phi(X\beta) = \int_{-\infty}^{X\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

This ensures that the predicted probabilities are limited between 0 and 1. We estimate the model via maximum likelihood estimation methods, where

$$\begin{aligned} L &= \Pr(Y_1 = y_1, Y_2 = y_2, \dots, Y_N = y_N) \\ &= \prod_{Y_i=0} [1 - F(X_i\beta)] \prod_{Y_i=1} F(X_i\beta) \\ &= \prod_{i=1}^N [1 - F(X_i\beta)]^{1-y_i} F(X_i\beta)^{y_i} \end{aligned}$$

is the likelihood function. The log-likelihood function that has to be maximised is then given by:

$$\ln L = \sum_{i=1}^N \{(1 - y_i) \ln[1 - F(X_i\beta)] + y_i \ln F(X_i\beta)\}$$

This maximum likelihood estimator is consistent, asymptotically efficient, and asymptotically normal distributed.

However, the estimated coefficients do not quantify the influence of the explanatory variables on the probability of being overeducated. The estimated coefficients are rather parameters of the latent, not observable model. Therefore, we calculate marginal effects as:

$$\frac{\partial P(y_i = 1 | x_i)}{\partial x_i} = \frac{\partial E(y_i | x_i)}{\partial x_i} = \varphi(x_i\beta)\beta$$

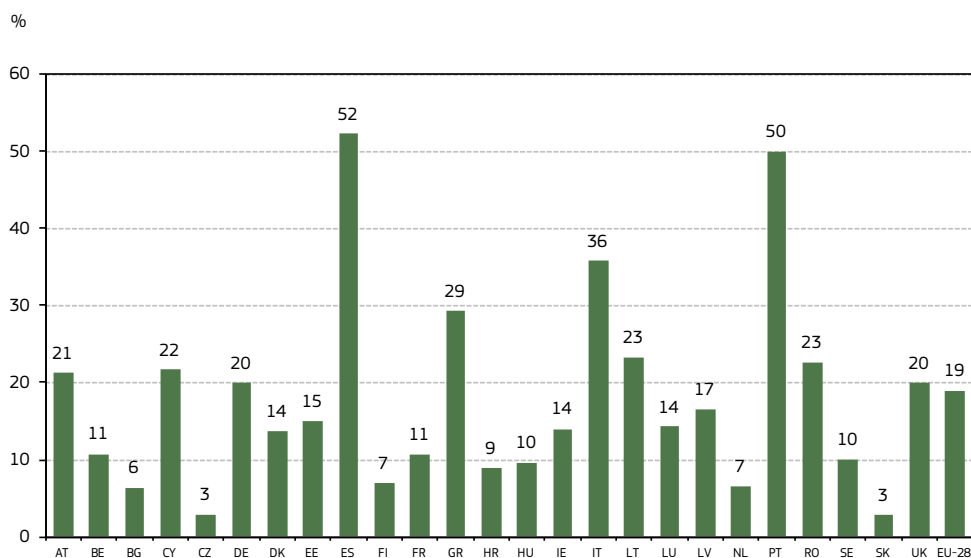
This marginal effect of any explanatory variable x_i is the effect of one unit change of this variable on the probability of being overeducated, given that all other explanatory variables are constant. Furthermore, these marginal effects depend on x_i , so we need to estimate the marginal effects at a specific value of x_i . We report marginal effects at the means of the independent variables.

5. Results

5.1 Descriptive results

Figure 1 depicts the incidence of overeducation for medium-skilled workers in 2013. As can be seen, the share of overeducated workers varies considerably between countries in both skill groups. While about half of the medium-skilled workers are overeducated in Spain (ES) and Portugal (PT), this holds for only 3% in the Czech Republic (CZ) and Slovakia (SK). With respect to the geographical distribution, it is a striking pattern that the four Southern European countries most severely affected by the current sovereign debt crisis (Greece, Italy, Portugal, Spain) all exhibit rates well above the EU average of 19%. Apparently, one reflection of the crisis in these countries is also a high risk of skill mismatches for medium-skilled workers. Potentially, this is an outcome of the general downturn of local labour demand, forcing workers to accept inadequate positions to avoid unemployment. At the same time, overeducation rates for medium-skilled in the Middle and East European countries that joined the EU in 2004 all rank clearly below the EU average. This regional particularity is confirmed by a previous report (European Commission, 2012), albeit not differentiating by educational level. The result could be related to the ongoing wave of labour migration from the Eastern to the Western part of the EU. Workers facing a scarcity of job offers adequate to their qualifications on their national labour markets prefer to migrate abroad instead of working in low-qualified jobs at home. This gains support by the fact that overeducation is observed to be especially rare in those countries exhibiting a common border with Germany and/or Austria.

Figure 1: Percentage of overeducated on all medium-skilled employees, 2013

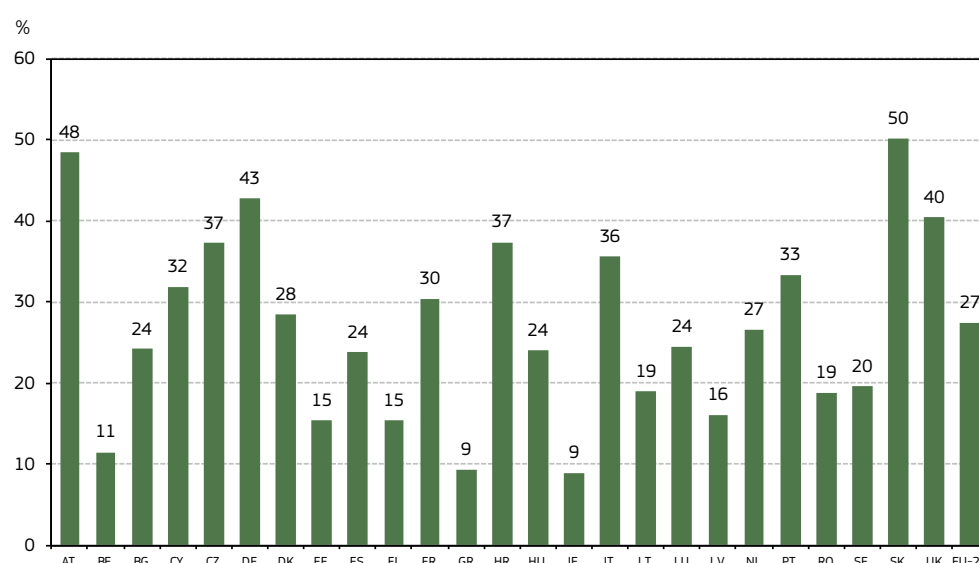


EU-28: without Malta, Poland and Slovenia

Sources: EU-LFS (2013), HWWI (2015).

For high-skilled workers, the picture looks quite different (see Figure 2). The overall incidence of overeducation is considerably more widespread in the majority of countries. The only exceptions to this are Spain, Greece, Ireland, Lithuania, Latvia, Portugal and Romania, where the overeducation risk of the medium skilled is still higher. The average percentage of overeducated workers on all the highly skilled is 27%. The geographical pattern also differs decisively. The Southern countries do not stand out with particularly high rates. The rate in Greece is with 9% (together with Ireland) even measured to be the lowest. Again, a combination of migration and unemployment could serve as an explanation for these numbers. The exodus of well-trained young graduates from the crisis countries (Verwiebe et al., 2014) has created a scenario where the remaining population of high-skilled is either well matched (and therefore has no emigration incentive) or unemployed, generating the statistical result of a seemingly high matching efficiency. For example, 52% of migrants with Spanish, Italian or Greek nationality who have moved within the last three years before 2011 were highly skilled, compared to 28% of nationals of the NMS-8, where by contrast 49% of migrants were medium educated (Boll et al., 2014c). Nevertheless, it is astonishing that in economically successful countries like Germany and the UK, educational mismatches among university graduates are apparently a much more frequent phenomenon. Confronting these results with the widespread fear of skill bottlenecks in these countries, the picture of a serious local mismatch of jobs and qualifications emerges. In the end, in identifying its roots, everything points to the educational systems as prime determinants of the skill composition. Moreover, contrary to our analysis for the medium-skilled, it is notable that countries like the Czech Republic and Hungary also exhibit rates above average, indicating that the emigration channel has so far been unable to clear an excess supply of human capital in the East European countries. In all, the EU-internal country differences in educational mismatch show that the European community is still far away from reaping the full benefits from cross-border labour migration created by the common market legislation.

Figure 2: Percentage of overeducated on all high-skilled employees, 2013



EU-28: without Malta, Poland and Slovenia

Sources: EU-LFS (2013), HWWI (2015).

In addition to the distribution by countries, the distribution by worker characteristics offers additional insights. Table 1 provides information on the percentage of overeducated and not overeducated high- and medium-skilled workers by sex, age, and field of study in the aggregate cross-country sample. As can be seen, highly-skilled men are slightly more frequently overeducated (29.3%) than women (25.7%) in this education group. In the group of medium-skilled-workers, the frequency of being overeducated hardly differs by gender. This does not need to imply that differences in match quality among skill groups are lower for women. It can also hint at gender differences in worker characteristics and the decision to work at all, channels whose roles will be investigated later on.

Table 1: Percentage of overeducated and not overeducated persons by sex, age, and educational attainment

	High-skilled		Medium-skilled	
	Overeducated	Not overeducated	Overeducated	Not overeducated
Sex (in % on all employees)				
- Male	29.3	70.7	13.2	86.8
- Female	25.7	74.3	13.1	86.9
Age (in % on all employees)				
- 15–24	42.8	57.2	15.7	84.3
- 25–34	28.6	71.4	16.1	83.9
- 35–44	26.9	73.1	13.9	86.1
- 45–54	26.5	73.5	11.6	88.4
- 55–64	25.1	74.9	10.2	89.8
- 65–74	32.4	67.6	15.3	84.7
Field of study (in % on all employees)				
- Teaching, education	15.4	82.4	20.5	79.5
- Humanities, languages, arts	27.2	72.8	11.7	88.3
- Social sciences, business and law	28.7	71.3	12.3	87.7
- Science, mathematics, computing	27.3	72.7	14.3	85.7
- Engineering, manufacturing, construction	31.2	68.8	12.2	87.8
- Agriculture, veterinary	46.5	53.5	19.3	80.7
- Health, welfare	21.4	78.6	8.3	91.7
- Services	40.9	59.1	14	86

Sources: EU-LFS (2013), HWWI (2015).

Moreover, among high-skilled workers the youngest (15–24 years) and the oldest group (65–74 years) of high-skilled workers are more frequently overeducated than workers from other age groups. A lack of actual work experience and a missing occupational focus of study programmes are potential factors to explain why young high-skilled workers are at a particular risk. For old high-skilled workers, other in-

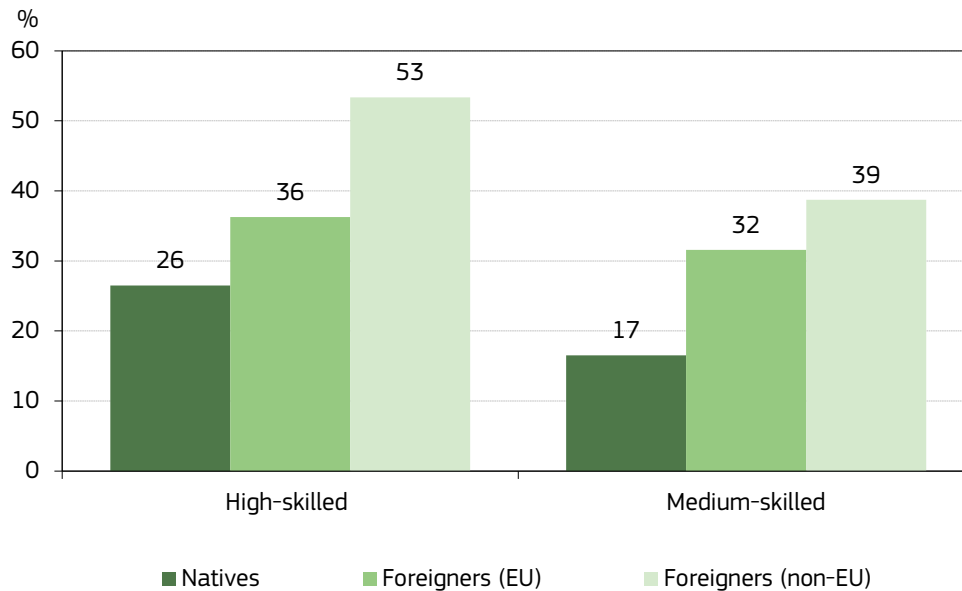
fluences should dominate, potentially related to formally low-ranked but yet re-deeming consulting activities executed in retirement. Given the voluntary nature of those activities, a welfare judgement of overeducation patterns should always discriminate carefully between age groups. Nevertheless, it has to be stressed that both age groups lie outside the core segment of high-skilled workers and are therefore rather thinly populated in our sample, limiting their usefulness for general conclusions. Again, the picture looks quite different if we look at medium-skilled workers. This group of workers is not only less likely to be overeducated in general, but the difference between different age groups is also less marked. Intuitively, this could both be ascribed to a stronger occupational focus of educational programmes at the medium level, facilitating young workers' access to adequate positions, and to a high relevance of practical work experience, helping older workers to maintain qualified jobs. Again, this result requires a more thorough analysis in the following section.

The comparison by field of study shows that high-skilled workers from the fields agriculture, veterinary and services are much more frequently overeducated than high-skilled workers from the fields teaching, education, and health and welfare. This becomes intuitive by considering the narrow occupational focus of the latter fields, thoroughly preparing graduates to become teaching and health professionals, respectively. Indeed, in our sample, 69% of all high-skilled graduates from teaching and education worked as teaching professionals, while 67% of high-skilled graduates from health and welfare worked as health (or health associate) professionals. Graduation has the role of an entry ticket into these occupations. The result is a high degree of internal educational homogeneity (Abraham et al., 2011). Existing research on social closure shows that this restriction practice is associated with positive wage effects, further stabilising the field-occupation linkages (Weeden, 2002). On the other hand, in study programmes in the field of services, the occupational focus tends to be much wider. This is reflected by a high variation in the occupations chosen by graduates from services in our sample. As a consequence, entry barriers into service-related occupations are largely weak, implying insufficient protection from falling back to low-qualified jobs. Consequently, high-skilled graduates from services rank second-highest in terms of overeducation frequency. However, among medium-skilled workers, the picture looks a little different. Here, the field teaching and education is associated with the highest overeducation share. To some extent, such shifts could be attributed to the skill-dependency of occupational profiles associated with formally the same field, limiting comparability of fields across skill levels. For instance, medium-skilled graduates from teaching and education only work to 48% as teaching professionals, while another significant share of 15% sorts into the occupational group of legal, cultural and social associated professionals.

Furthermore, the incidence of educational mismatches could also vary between domestic and foreign workers (as differentiated by nationality). To the extent that foreign workers have immigrated, they might find it difficult to align their human capital accumulated abroad with the conditions on the local labour markets. Language and cultural barriers could also block their access to formally adequate positions. In analysing the role of nationality, it is meaningful to additionally distinguish between EU- and non-EU foreigners in our sample, as the latter group faces additional legal restrictions (e.g. requirement of work permit). Indeed, Figure 3 documents a corresponding variation in overeducation frequencies across the three groups. Among high-skilled non-EU foreigners, even a slight majority of workers is assessed to be mismatched. Among native high-skilled, this is only observed for about one fourth. EU foreigners are located, as expected, somewhere in the middle. Similar relative

discrepancies occur in the group of medium-skilled workers. At the same time, the fact that overeducation is more frequent among the high-skilled holds for natives and foreigners alike.

Figure 3: Share of overeducated workers by skill level and worker origin (cross-country sample), 2013



Sources: EU-LFS (2013), HWWI (2015).

Finally, more detailed knowledge on the roots of overeducation can be gained by comparing groups based on a nesting of factors like age, gender and nationality. This could prove insightful especially with respect to the so far unexplained gender differences. In the causal analysis performed in the next section, it will be examined in how far the relationship between overeducation risk and determinants as age and field of study differs between male and female workers.

5.2 Estimation results

Table 2 presents estimated coefficients of the aggregate cross-country model for high-skilled and medium-skilled workers separately (for an introduction to the interpretation of regression results, see Methodological Guides). Estimation results for single countries are listed in Table A 1 to A 10 in the Appendix. The following discussion focuses on results from the cross-country model. Results for single countries are only reported if they clearly deviate from the aggregate ones.

Table 2: Estimation results (cross-country estimation)

	High-skilled		Medium-skilled	
	<i>Coeff</i>	<i>SE</i>	<i>Coeff</i>	<i>SE</i>
Personal characteristics				
Female	-0.008	0.013	-0.004	0.008
Age groups (reference: 25–34 years):				
- 15–24 years	0.140***	0.022	-0.008	0.007
- 35–44 years	0.011	0.009	-0.009**	0.004
- 45–54 years	-0.006	0.009	-0.029***	0.004
- 55–64 years	-0.007	0.010	-0.026***	0.005
- 65–74 years	0.060**	0.024	-0.010	0.012
Married	-0.029***	0.007	-0.004	0.003
Foreigner: EU countries	0.078***	0.014	0.112***	0.010
Foreigner: non-EU countries	0.167***	0.016	0.085***	0.008
Field of study (reference: social sciences, business and law):				
- Teaching, education	-0.102***	0.010	0.089***	0.019
- Humanities, languages, arts	0.012	0.012	-0.018*	0.009
- Science, mathematics, computing	-0.008	0.010	-0.017**	0.008
- Engineering, manufacturing, construction	-0.011	0.007	-0.058***	0.004
- Agriculture, veterinary	0.094***	0.017	-0.024***	0.006
- Health, welfare	-0.096***	0.010	0.038***	0.009
- Services	0.108***	0.015	-0.041***	0.005
Household characteristics				
Number of unemployed adults ¹	0.048***	0.007	0.017***	0.004
Number of inactive persons ¹	0.013***	0.003	0.001	0.002
Persons aged 75 or older ¹	0.057	0.118	-0.036	0.049
Number of children between 0 and 5 years ¹	-0.006	0.005	0.002	0.003
Number of children between 6 and 11 years ¹	-0.008	0.005	0.000	0.003
Number of children between 12 and 14 years ¹	-0.001	0.008	0.001	0.004
Job characteristics				
Firm size (reference: < 10 persons):				
- 11–19 persons	-0.014**	0.007	0.005	0.004
- 20–49 persons	-0.054***	0.006	-0.005	0.003
- 50 and more persons	-0.053***	0.006	0.015***	0.003
Marginal employment ²	-0.019	0.015	-0.014*	0.007
Temporary contract	0.038***	0.007	0.027***	0.004
Usual working hours	-0.058***	0.010	-0.001	0.006

Usual working hours squared	0.001	0.001	0.000	0.001
Tenure	-0.002	0.006	-0.006**	0.003
Tenure squared	-0.006***	0.002	-0.004***	0.001
Participation in LLL	-0.036***	0.006	0.021***	0.004
Second job	0.007	0.008	0.013**	0.006
Degree of urbanisation	0.009***	0.003	-0.005***	0.002
Interaction terms:				
Sex and married	0.021**	0.009	-0.009*	0.005
Sex and urbanisation	0.014***	0.005	-0.002	0.003
Sex and elder household members	0.036	0.131	-0.035	0.061
Sex and children:				
- number of children (0–5 years)	-0.030***	0.007	-0.009**	0.005
- number of children (6–11 years)	0.005	0.007	0.000	0.004
- number of children (12–14 years)	0.010	0.011	0.003	0.006
Sex and age groups (references: 25–34 years):				
- 15–24 years	-0.013	0.024	-0.023***	0.009
- 35–44 years	-0.013	0.011	-0.011*	0.006
- 45–54 years	0.009	0.011	-0.012**	0.006
- 55–64 years	0.018	0.013	-0.024***	0.007
- 65–74 years	-0.008	0.034	-0.049***	0.014
Sex and field of study (reference: social sciences, business and law):				
- Teaching, education	0.017	0.013	0.044**	0.018
- Humanities, languages, arts	-0.021	0.015	0.017	0.013
- Science, mathematics, computing	-0.014	0.015	0.023	0.015
- Engineering, manufacturing, construction	-0.026**	0.011	0.017***	0.006
- Agriculture, veterinary	-0.036	0.022	0.043***	0.013
- Health, welfare	0.014	0.012	0.012	0.008
- Services	0.010	0.021	0.013*	0.008
Observations	248,230		431,542	

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry and country included. LLL: Life long learning. Coeff: coefficients, SE: robust standard errors.

Sources: EU-LFS (2013), HWWI (2015)

Individual characteristics

In the cross-country perspective, the base level of gender proves to be insignificant for both medium and high-skilled workers. However, this does not imply the general absence of a gender effect on overeducation. This is shown by the interaction terms of gender with other covariates, precisely with household composition variables and degree of agglomeration. The insignificant base term of gender solely implies that there exists no significant gender difference in overeducation risk for the subgroup of unmarried workers living in urban areas in households with neither people older than 75 nor children younger than 15. The partial significance of the interaction terms shows this is different for other subgroups. For instance, the interaction of gender with marital status is significant for members of both skill groups, albeit with opposite signs. Additionally, the base level of marital status is significantly ne-

gative for high-skilled workers and insignificant for medium-skilled workers. Figure 4 displays the composed marginal effects of being member of a certain subgroup.

Figure 4: Marginal effects of the interaction of gender and marital status (reference: unmarried men)



Sources: EU-LFS (2013), HWWI (2015).

Hence, among high-skilled workers, married men face a significantly lower overeducation probability (about 2.8% at the median) than unmarried men. At the same time, the probability is also clearly lower than for married women, which is in line with Frank's theory of differential overqualification (1978). Among medium-skilled workers, married and unmarried men do not exhibit significantly different probabilities, but they are significantly higher than those for married women. The Frank theory thus does not find confirmation by the aggregate sample of medium-skilled workers.

Further heterogeneity comes into play when comparing results for single countries. Regarding the base term, Italy stands out as the only large country that deviates from the aggregate results in both skill segments. The coefficient is measured to be significantly positive for high-skilled and negative for medium-skilled workers, implying that in Italy even the job perspectives of unmarried childless women differ from their equally skilled male counterparts. With the interaction term, the UK represents an outlier with its positive coefficient for medium-skilled workers, showing that being married raises the overeducation risk for both high- and medium-skilled women in this country.

With regard to the interplay of gender and urbanisation, results likewise tend to differ between high-skilled and medium-skilled workers. Among high-skilled workers, women living in rural areas face a higher overeducation risk than women in urban areas in the cross-country estimations, an outcome again coinciding with the views of Frank. Another interpretation is that in rural areas women are significantly more at risk than men. For medium-skilled workers, this interaction term is insignificant, thus not pointing at a spatial dimension of the gender effect. At first glance, this outcome seems surprising, as high-skilled workers are commonly observed to be more mobile than workers with lower qualification levels (Greenwood, 1975). Therefore, one could expect their job quality to be less affected by conditions in their area of living. At the same time, however, theories of knowledge spillovers suggest that high-skilled workers benefit to the largest extent from working in urban areas (Du-

ranton and Puga, 2004). Within rural areas, high-skilled women are thus expected to be more severely impaired by these spatial restrictions in their job search than medium-skilled women. This idea is also in line with the coefficients of the base level of urbanisation, which are positive for high-skilled workers and negative for low-skilled workers.

In the single-country estimations, the significance of the interaction of urbanisation and gender in the high-skilled segment is not observed for all large countries, for instance not for France and the UK. For these two countries, no rural-urban gap in the overeducation probability of women can be proved. This could hint at a higher degree of worker mobility and/or at a stronger influence of the female partner on a couple's co-location decision in these countries. Furthermore, the insignificance of the same term for medium-skilled workers in the aggregate estimation is revealed to be the outcome of contrary effects at country level. In countries like Hungary and Romania, medium educated women are at a significantly higher risk in rural than in urban areas, while the opposite holds in countries like Greece and Sweden. The interplay of space and gender in determining the probability of overeducation thus sketches a more clear-cut pattern for the highly educated than for the medium educated.

Of less complex nature is apparently the link to nationality. For both high- and medium-skilled workers, the overeducation risk is measurably higher for foreigners. We also tested an alternative model in which nationality is also interacted with gender. However, results proved insignificant for both skill levels, suggesting that the impact of nationality is similar for male and female workers. To the extent that we can equate foreigners with 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. Some interesting discrepancy however emerges in the comparison of EU- and non-EU foreigners. Among high-skilled workers, the overeducation probability is clearly lower for EU- than for non-EU foreigners. Among medium-skilled workers, however, this is not observed. Again, this is a slightly surprising result. One could expect that for high-skilled workers distance matters less, as they should be better trained in aligning their skills with changing local conditions. Thinking in another direction, an explanation might be the existence of general legal or social barriers non-Europeans face in accessing high-skilled jobs in the EU. This fact would obviously be of less relevance for the overeducation risk of medium-skilled workers. Results at country level mostly fit this overall picture, apart from a few outliers. For instance, no significant differences between native and foreign medium educated workers are identified in the Netherlands. Among the highly educated, EU foreigners are even predicted to be less exposed to the risk of overeducation than Dutch workers. Part of these country differences could be related to cultural ties, to the existence of migrant networks as well as to the particularities of local labour markets.

For high-skilled workers, cross-country results on the role of worker's age seem to support the U-shaped relationship found by Green and McIntosh (2007) as well as Joona et al. (2012). Compared to the reference category 25–34 years, both the youngest group of 15–24 years old and the oldest group of 65–74 years old workers are predicted to be at a significantly higher risk. As the interaction terms of gender and age remain insignificant, this pattern holds for male and female high-skilled workers alike. Given that we control for marginal employment, it cannot simply be dismissed as reflecting age-specific spare-time work like student jobs or jobs for retirees. Rather, two conclusions seem to emerge. First, the high risk faced by the youngest group points at the existence of significant entry barriers that young

high-skilled workers face when accessing the labour market immediately after graduation. Given the increasing supply of young graduates from university bachelor programmes as part of the Bologna process, the outcome is worrying in its potential implications for the future career paths of these cohorts. Second, it is interesting to see that the category of 56–64 years old workers, which in many countries represents the group immediately before the regular retirement age, is not yet affected by the U-shaped turn. Skill depreciation thus does not seem to be a general concern for high-skilled workers, at least not with respect to the adequacy of jobs. The measured risk increase for high-skilled workers older than 65 years should not be considered a cause for worry here, as it mainly concerns voluntary activities (e.g. in the area of consulting and supervision) of persons who could afford to stay outside the labour market.

In contrast, cross-country results for medium-skilled workers reveal a startlingly different picture, both concerning the overall distribution among age groups and the nature of gender differences. For male workers, it is the reference group of 25–34 years old that is measured to exhibit the highest probability of overeducation, with probabilities for all other age groups being significantly lower. Hence, among male workers with medium education the youngest group is not exposed to a particularly high risk, which might be explained by the shorter training period and the stronger occupational focus compared to higher educational levels. At the same time, results for the oldest cohorts point at a beneficial role of work experience. Coefficients for the three oldest age groups are clearly smaller than for the two younger groups, indicating a reduced overeducation risk for male workers older than 34. For medium-skilled workers, this beneficial role seems to be of even higher importance than for high-skilled workers, possibly due to the lower requirements concerning formal education. However, at least for the group of 55–64 years old the statistical effect of non-random labour market exit (outselection) has to be kept in mind: workers trapped in bad matches (especially those with physically demanding work) have higher incentives to retire prematurely. Coefficients of the interaction terms of age and gender are throughout negative and significant. This implies that the risk-reducing effect of age (starting from the group of 35–44 years old) is not only maintained for women, its magnitude is even larger than in the case of men. As for male workers, this result should be interpreted against the background of selection effects in the context of labour market exit, which are likely to be stronger and to occur earlier in life for female workers.

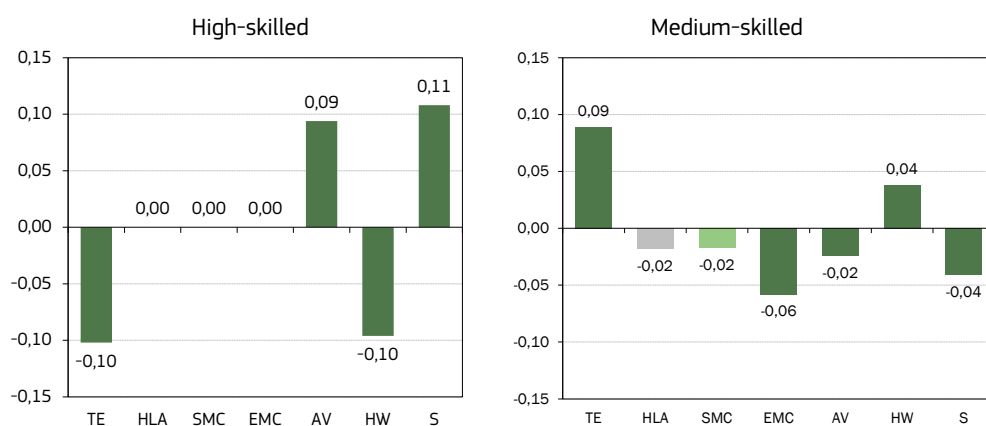
In all, results for the highly skilled are in the first half of the age distribution, broadly compatible with both career mobility theory and human capital theory (see above), and in the second half with ideas of human capital depreciation. For the medium skilled, it is only the second half of the age distribution which is in line with major theories, again both with career mobility theory and human capital theory.

As to be expected, results at country level show some degree of heterogeneity. The European-wide patterns of age coefficients are for both high- and medium-skilled workers, nevertheless largely confirmed at country level. For the high-skilled workers, Italy constitutes an outlier among the large countries. Here, overeducation risk is continuously declining with increasing age for the high-skilled. Among medium-skilled workers, Greece is a special case with a particularly high overeducation probability for the youngest group of 15–24 years old workers.

Finally, our results for the impact of field of study deserve some attention. In this regard, the comparison to Ortiz and Kucel (2008) is most informative, as they apply the same categorisation of fields and discriminate between tertiary and non-

tertiary graduates in their estimation. However, in contrast to them, we added a differentiation of effects by gender, which proves revealing. First, among male high-skilled workers, the cross-country estimation yields the highest overeducation risk for graduates from the field services, which is consistent with the results of Ortiz and Kucel (2008). The difference in risk to the reference group social sciences, business and law is highly significant (see Figure 5). The same holds for the group with the second largest risk in our data, students of agriculture and veterinary medicine. A deviation from Ortiz and Kucel (2008) emerges with respect to students of humanities. In contrast to their results this group does not, in our estimation, exhibit a significantly higher risk than the reference group. Another difference concerns students of teaching and education, which in our approach enjoy the smallest overeducation probability, also significantly smaller than the reference group. As explained in the previous section, this can be rationalised by the phenomenon of social closure. Graduates from teaching have overcome the hurdle for working as a teaching professional. They therefore enjoy access to a number of adequate positions with limited competition. Due to the wider occupational focus of their studies, graduates from many other fields are much less protected, therefore facing a higher risk of ending up in mismatches.

Figure 5: Marginal effects of field of study for male workers (cross-country estimation), 2013



TE: Teaching, Education. HLA: Humanities, Languages, Arts SMC: Science, Mathematics, Computing. EMC: Engineering, Manufacturing, Construction. AV: Agriculture, Veterinary. HW: Health, Welfare. S: Services. Reference group: Social Sciences, Business and Law.

Dark green: significant at 1 %-level. Light green: significant at the 5 % level. Grey: significant at the 10 % level.

Sources: EU-LFS (2013), HWWI (2015).

In analysing the interaction with gender, the insignificance of most interaction terms suggests that the relationship between overeducation risk and field choice is qualitatively very similar for male and female graduates from tertiary education. The prime difference concerns the field health and welfare, where the risk-reducing effect compared to the reference group of social sciences, business and law is estimated to be even more pronounced for women. This could be due to gender differences in field selection within the rather broad reference group. For instance, a higher popularity of social sciences with their rather vague job profiles among female students

could raise the overall risk of the reference group. However, the significance of this effect remains weak.

For medium-skilled workers, results of our estimation are almost turned upside down: male graduates from teaching and education are at the highest risk, graduates from services at the second lowest (after engineering and construction). As explained above, the discrepancy most likely results from differences in the matching of educational fields and occupational groups across educational levels. These limit comparability across skill groups. Here, the interaction with gender reveals a particularly high risk for female medium-skilled workers in the field of health and welfare. It is likely that occupational segregation plays a major role in this, given the female image of low-qualified jobs in the area of caretaking and medical assistance. Segregation could also help to explain the observed gender difference in teaching and education, where again the high relative risk of medium-skilled men is even exceeds that for women. In contrast to results for the high-skilled, the gender bias in field-specific risks is estimated to be highly significant. Hence, among medium-skilled workers, the sorting into certain fields of study seems to have a stronger impact on overeducation risk for women than for men.

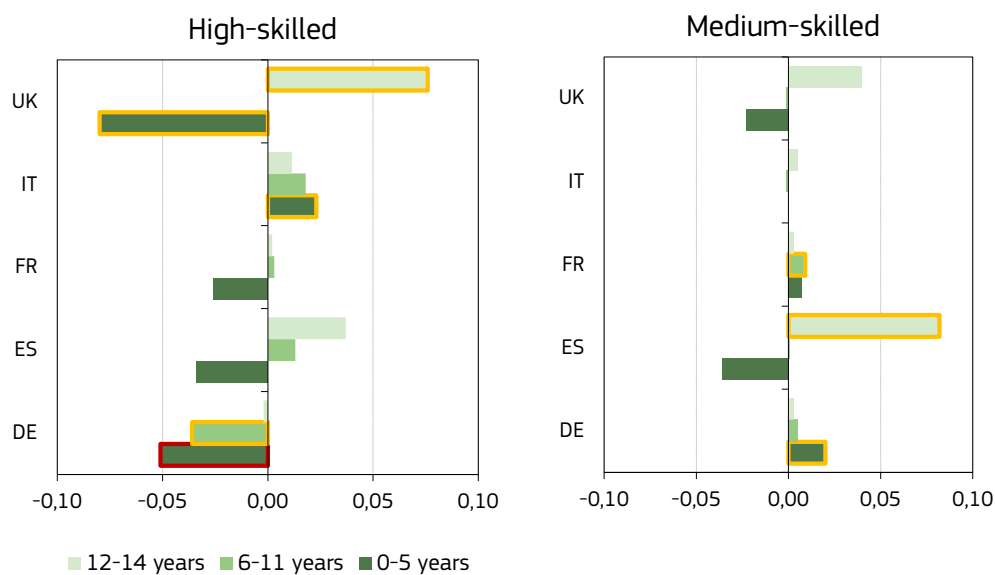
At the country level, the single estimations show the highest degree of conformity with respect to the situation of high-skilled graduates from services. Among the larger countries, only in Greece and Italy do these graduates not show a significantly higher overeducation probability than the reference category. The low probability for high-skilled graduates from teaching is also remarkably widespread. In the high-skill segment, the strongest diversity of coefficients across countries can be observed for the fields humanities and engineering. Regarding humanities, for instance in Italy and Portugal a significantly positive coefficient for the base term is estimated (again implying a higher overeducation risk than for the reference group), while it is negative in Germany. Results for engineering are even more equally divided. Comparing field effects for medium-skilled workers, it is again humanities where results are particularly mixed. The same can be said about medium-skilled graduates from services. The overall picture can thus be described as follows: while some combinations of field and skill level are systematically associated with higher or lower overeducation probabilities throughout Europe, the role of others is highly country-specific.

Household characteristics

A salient feature of our estimates for the household-specific determinants is the gender dichotomy. Starting with the role of children in the household, the coefficients of the base terms are generally insignificant. This means in our set-up that the overeducation risk of male workers is not affected by children of any number and age composition. For female workers, the interaction terms with gender are relevant. Here, the result seems to hinge upon the age of children as well as the worker's educational level. For high-skilled female workers, having an additional child below the age of six is predicted to reduce the overeducation probability significantly. This would match our reservation wage argument from above: the reward should be high enough to cover the (monetary and emotional) costs of placing the child into external care. In turn, high rewards are typically associated with high skill requirements. For older children, the association remains insignificant, which is also in line with intuition, given that school-age children require less intense care.

For medium-skilled workers, the risk-reducing effect of small children is of lower magnitude and only weakly significant. This does not necessarily imply that medium-skilled females face lower costs of childcare. It could also result from the generally lower overeducation risk for medium-skilled workers. Another reason might be that medium-skilled workers are on average expected to be less wealthy than the high-skilled, which could force them to also accept barely adequate jobs when living with children. In the estimations undertaken at country level, having children below the age of six cannot be proved to raise the overeducation risk of high-skilled workers for the majority of countries. Nevertheless, we observe a serious degree of heterogeneity also among large countries (see Figure 6).

Figure 6: Marginal effects of no. of children for female workers in five West European countries, 2013



Yellow frame: significant at the 5 % level. Red frame: significant at the 1 % level.

Sources: EU-LFS (2013), HWWI (2015).

Furthermore, the presence of unemployed persons in the same household is estimated to increase the likelihood of overeducation for both skill classes. This is interesting, not least because it is a so far a new result in the overeducation literature. One interpretation could be that the need to financially support unemployed household members induces workers to avoid their own unemployment by accepting even comparatively bad matches. This result is confirmed for many, if not for all countries in the separate estimations. Exceptions for which the coefficients show a reverse sign are only found at the medium-skill level and comprise two countries, Germany and Lithuania.

When comparing the role of the unemployed with those of inactive household members, the influence of the inactive is measurably smaller in the cross-country estimations, as expected. For medium-skilled workers, the influence is insignificant. This seems to point at a linkage between household composition and job-related

productivity: workers living together with the unemployed might themselves be less productive on average, a fact that reduces their chances of finding an adequate match for their formal education. In other words, for the medium educated the correlation points rather to selection than to causality. At the country level, this distinction is largely confirmed. In a clear majority of countries, the effect of unemployed household members exceeds the effect of the inactive. A notable exception is marked by high-skilled workers in Italy, for whom only inactive household members significantly contribute to a higher overeducation risk. Again, Germany represents an outlier for medium-skilled workers: just like in case of unemployed household members, inactive members are predicted to reduce the overeducation probability, even though to a lower degree.

By contrast, the presence of persons in the household at the age of 75 and above is not estimated to change the risk of overeducation in the cross-country analysis. Among the few countries for which sufficient data are available, one exception marked again is Germany, where the presence of the elderly is, at least for high-skilled workers, associated with a higher overeducation probability. The other exception is Hungary, in the sense that the association is measured to be significantly negative for medium-skilled workers.

Job-related characteristics

Regarding job characteristics, results are partially surprising. First and foremost, this concerns the role of marginal employment.⁵ While being insignificant for high-skilled workers, the coefficient is weakly significant and negative for medium-skilled workers. This means that marginally employed workers are at a lower risk of becoming overeducated, which contradicts the intuition outlined above. Even more surprising, in the estimations at country level this result is confirmed for a clear majority of countries both in the segments of high- and medium-skilled workers. For instance, the negative coefficients are even highly significant for high-skilled workers in France and the UK. Prominent exceptions are Germany and Italy, where marginal employment is linked to a higher overeducation risk for both skill segments. Descriptive analysis reveals this difference results from the fact that marginally employed workers tend to select into different occupational groups in these countries. For instance, managers and professionals make up a considerably larger share among marginally employed workers in the UK (48.4%) than in Germany (39.3%). This might reflect country differences in the social perception of certain jobs. However, we need to be cautious with our interpretation, given that we only define marginal employment based on information on working hours.

In contrast, the coefficients for working in a temporary position are clearly positive for high- and medium-skilled workers, which is in line with expectation. Again, one explanation could be that the transitory nature of fixed-time jobs could convince people to accept less ideal matches. Contrary results at country level are rare, not including larger countries except for the Netherlands for medium-skilled workers. Having a second job does, in the cross-country estimations, only raise the overeducation probability at a medium-skill level. For high-skilled workers, a relationship cannot be confirmed. Special cases here are Ireland and the UK, where the effect is estimated to be positive for both skill types.

⁵ According to the legal definition: employment for which the wage is regularly not exceeding Euro 400 per month and / or employment which – during a calendar year – is restricted to two months or 50 working days (irrespective of the earnings).

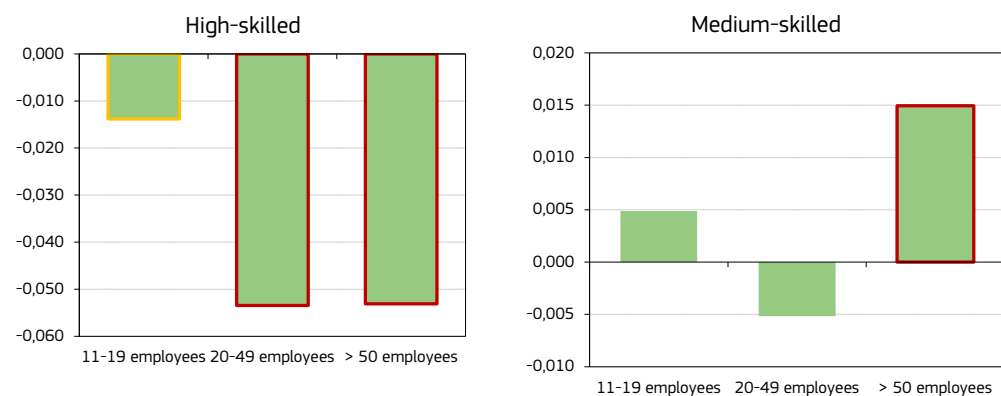
The role of training activities exhibits contrary signs for the high- and the medium-skilled in the cross-country estimations. For high-skilled workers, recent participation in LLL (life long learning) is associated with significantly lower overeducation risk for high-skilled, but significantly higher overeducation risk for medium-skilled workers. The risk reduction for high-skilled workers is also observed for almost all countries in the country-specific measurement. This indicates that the starting points for participating in training activities tend to be different for high- and medium-skilled workers. Among the highly educated, training priority lies within already favourable matches. Here, the view seems to dominate that only adequate jobs offer promising opportunities for career advancement through training. By contrast, training incentives for the medium educated are predicted to be higher under unfavourable matches. Here, the upgrading argument seems to apply, i.e. workers seek to escape inadequate positions by improving their human capital through training. This dichotomy receives some support by research on segmented labour markets. It has been documented that among workers on secondary labour markets (i.e. markets for high-turnover and low-pay work), both the incidence of overeducation and training participation positively affect career mobility, an effect that was estimated to be stronger than on primary markets (Dekker and de Griep, 2002). To the extent that medium-skilled workers are more frequently active in the secondary labour market, this can help to explain the especially high training incentives for the badly matched medium-skilled. While the negative coefficient of LLL participation for the highly skilled proves to be negative for the majority of countries, there exists considerable country variation in the coefficients for the medium-skilled. For instance, in Greece and Italy it is positive and thus opposed to the aggregate effect. In all, the outcome documents the ambiguous relationship between training participation and the incidence of skill mismatches outlined above.

Results for the remaining job-related characteristics show a slightly more uniform picture. Longer job tenure is associated with a shrinking risk of overeducation for members of both skill segments, which corresponds to our intuition discussed above. Workers in a skill mismatch are unlikely to achieve high levels of job satisfaction and are therefore not expected to remain in the current match for a long time. As the quadratic term is negative and highly significant, the risk reduction associated with any additional year is predicted to become even stronger with increasing tenure. For very long tenure, an overeducation status is thus especially rare. Moreover, while the negative association with overeducation risk cannot be statistically proved for all single countries, it is nowhere estimated to be positive except for medium-skilled workers in Romania. This can be interpreted as a high degree of stability, especially compared to the outcomes for other covariates. Less straightforward are the results for hours of work. In the aggregate approach, the expected negative coefficient can only be confirmed for the high-skilled. This might have something to do with better promotion opportunities for high-skilled workers. These could create incentives to show high levels of engagement, i.e. to spend much time at the workplace, which is in turn rewarded through advancement into better positions. For medium-skilled workers, the corresponding incentives are lower, diluting a relationship between hours of work and overeducation risk. Moreover, the quadratic terms remain insignificant for both skill classes. Variation of the estimates among countries mostly concerns the quadratic term, reaching from a clearly positive to a clearly negative influence. This suggests an overlapping of several contrary influences, which are hard to disentangle in the individual case.

Moreover, the estimated coefficients for firm size also show conflicting patterns (see Figure 7). For the high-skilled, the overeducation risk is predicted to decline

with increasing firm size. This result seems to fit the notion that larger firms can achieve a more precise matching of applicants and positions. Moreover, they offer more internal opportunities for advancement. For the medium-skilled, however, this is not observed. Workers in firms with more than 50 employees are at a significantly higher risk than those in very small firms. This observation is presumably technology-related, in the sense that it reflects a relationship between occupation and firm organisation: manual jobs in industrial mass production are typically executed within large organisations. For instance, for medium-skilled workers within the occupational group of plant and machine operators and assemblers, the share of workers employed in firms which are measured to have more than 50 employees is equal to 49.6% in our dataset, which clearly exceeds their share across occupations of 34.7%.

Figure 7: Marginal effects of firm size (cross-country estimation), 2013



Reference group: < 11 employees. Yellow frame: significant at the 5 % level. Red frame: significant at the 1 % level.

Sources: EU-LFS (2013), HWWI (2015).

Finally, we also undertook additional estimations including further explanatory factors at the regional level (NUTS 2), such as the regional unemployment rate and 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 for the EU-28 countries in a unified framework. Based on data from the European Labour Force Survey (EU-LFS), a series of individual, household, job-related and regional characteristics were used as explanatory factors in a probit model explaining the probability that a worker can be considered overeducated. Estimations were undertaken both for a cross-country sample and a set of 25 EU countries, selected based upon data availability. At each regional level, the sample was further split into two subsamples of high- and medium-skilled workers.

For most potential determinants, sign and magnitude of the estimated impact exhibits considerable variation both among countries and skill segments, justifying our disaggregated analysis. Results that are less scenario-sensitive are the higher overeducation risk of foreigners compared to native workers, the lower risk for persons with longer job tenure, the higher risk for persons in temporary compared to persons in permanent positions and the risk-increasing effect of the presence of unemployed and inactive household members. Others apparently specific to high-skilled workers are a U-shaped relationship between age and overeducation risk and the significant risk increase female workers face from having a child below the age of six. Moreover, the risk was shown to vary significantly with the chosen field of study even when controlling for all other measurable characteristics. Among high-skilled workers, graduates from the field of services exhibit the highest, teaching graduates the lowest EU-wide probability of overeducation. For medium-skilled workers, almost the inverse outcome was obtained, with graduates from teaching facing the highest and graduates from engineering facing the lowest overeducation probability. Regarding the more controversial results, of course, all of our results need to be interpreted against the background of potential distortions like omitted variable bias and two-sided causality, dangers that can hardly be avoided within a cross-sectional analysis even in a data-rich setting like ours.

Our results provide a stimulus for future research with respect to several aspects. First, the striking discrepancies observed in the role of many determinants between highly and medium educated workers clearly deserve some attention. It would be interesting to know to what extent these results reflect a genuine treatment of persons at certain educational levels and to what extent they merely disguise the selection effect of educational programmes. Second, regarding the effects of household composition, a further differentiation could yield additional insights. For instance, among the inactive household members, it would be beneficial to distinguish between permanently inactive ones (e.g. due to physical disability) and those who would be willing to work under changed conditions (hidden reserve). Under the assumption of rational behaviour, the search pressure imposed by inactive household members on their active housemates can be expected to be lower in the case of voluntary inactivity, implying a different relation to overeducation. Third, our cross-sectional set-up could be extended to a panel dataset in order to study the dynamics of overeducation. This would allow us to analyse the impact of previous employment histories, thereby gaining insights into the persistence of the pheno-

menon over the lifecycle. Fourth, an investigation of the interplay of the roles of gender and selection into educational fields would contribute to our understanding of the gender discrepancies. Finally, in order to stress the policy relevance of these results, empirical analysis should link the overeducation phenomenon to individual welfare. First and foremost, this concerns the association with the wage level. This is a research question we will address in another module of our report.

7. References – Module 1

Abraham, M., Damelang, A., Schulz, F. and für Arbeit, B. (2011): Wie strukturieren Berufe Arbeitsmarktprozesse? Eine institutionentheoretische Skizze. *LASER Discussion Paper*, 55.

Alba-Ramirez, A. (1993): Mismatch in the Spanish labor market: Overeducation? *The Journal of Human Resources* 28(2), 259–278.

Allen, J., van der Velden, R. (2001): Educational mismatches versus skill mismatches: Effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers* 53(3), 434–452.

Baert, S., Cockx, B. and Verhaest, D. (2012): Overeducation at the start of the career: Stepping stone or trap? *IZA Discussion Paper*, Series No. 6562, 1–40.

Battu, H., Belfield, C. R. and Sloane, P. J. (2000): How well can we measure graduate overeducation and its effects? *National Institute Economic Review* 171(1), 82–93.

Bauer, T. (2002): Educational mismatch and wages: A panel analysis. *Economics of Education Review* 21(3), 221–229.

Becker, G. S. (1964): *Human Capital and the Personal Distribution of Income: An Analytical Approach*. Michigan, MI: Institute of Public Administration.

Berlingieri, F. and Zierahn, U. (2014): Field of study, qualification mismatch, and wages: Does sorting matter? *ZEW-Centre for European Economic Research*, Discussion Paper No. 14-076.

Boll, C. and Leppin, J. (2014a): Overeducation among graduates: An overlooked facet of the gender pay gap? Evidence from East and West Germany. *SOEP papers on Multidisciplinary Panel Data Research*, No. 627.

Boll, C., Leppin, J. and Schömann, K. (2014b): Who is overeducated and why? Probit and dynamic mixed multinomial logit analysis of vertical mismatch in East and West Germany. *HWWI Research Paper 149*, forthcoming in *Education Economics*.

Boll, C., Leppin, J. and Quitzau, J. (2014c): *Is the Euro Boosting Mobility? Labour Mobility in Europe During the Crisis Years*. A study prepared jointly by Berenberg and HWWI, Hamburg.

Büchel, F. and Mertens, A. (2000): Overeducation, undereducation, and the theory of career mobility. *IZA Discussion Paper*, Series No. 195, 1–28.

Büchel, F. and Pollmann-Schult, M. (2001): Overeducation and skill endowments: The role of school achievement and vocational training quality. *IZA Discussion Paper*, Series No. 337, 1–28.

Büchel, F. and van Ham, M. (2002): Overeducation, regional labour markets and spatial flexibility. *IZA Discussion Paper*, Series No. 424, 2–13.

Büchel, F. and Battu, H. (2003): The theory of differential overqualification: Does it work? *Scottish Journal of Political Economy* 50(2), 1–16.

- Capsada-Munsech, Q. (2015): The role of social origin and field of study on graduates overeducation: the case of Italy. *Higher Education* 69(5), 779–807.
- Chevalier, A. (2003): Measuring over-education. *Economica*, 70(279), 509–531.
- Daly, M. C., Büchel, F. and Duncan, G. J. (2000): Premiums and penalties of surplus and deficit education. Evidence from the United States and Germany. *Economics of Education Review* 19(2), 169–178.
- Dekker, R. and de Griep, A. (2002): The effects of training and overeducation on career mobility in a segmented labor market. *International Journal of Manpower* 23(2), 106–125.
- Duncan, G. J. and Hoffmann, S. D. (1981): The incidence and wage effects of over-education. *Economics of Education Review* 1(1), 75–86.
- Duranton, G. and Puga, D. (2004): Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics* 4, 2063–2117.
- European Commission, Directorate-General for Employment, Social Affairs and Inclusion (2012): Employment and Social Developments in Europe 2012. Luxembourg: Publications Office of the European Union.
- European Union (2014): Quality report of the European Union Labour Force Survey 2013–2014 edition. Luxembourg: Publications Office of the European Union.
- Frank R (1978): Why women earn less: The theory and estimation of differential overqualification. *American Economic Review* 68(3), 360–373.
- Frenette, M. (2004): The overqualified Canadian graduate: The role of the academic program in the incidence, persistence, and economic returns to overqualification. *Economics of Education Review* 23(1), 29–45.
- Green, F. and McIntosh, S. (2007): Is there a genuine under-utilisation of skills amongst the over-qualified? *Applied Economics* 39(4), 427–439.
- Greenwood, M. J. (1975): Research on internal migration in the United States: A survey. *Journal of Economic Literature* 13(2), 397–433.
- Groot, W. (1996): The incidence of, and returns to overeducation in the UK. *Applied Economics* 28(10), 1345–1350.
- Groot, W. and Maassen van den Brink, H. (2003): The dynamics of skill mismatches in the Dutch labor market: Overeducation in Europe. *Current Issues in Theory and Policy*, 49–63.
- Joona, P.A., Gupta, N.D. and Wadensjö, E. (2012): Overeducation among immigrants in Sweden: Incidence, wage effects and state-dependence. *IZA Discussion Paper*, Series No. 6695, 1–39.
- Jovanovic, B. (1979): Job matching and the theory of turnover. *The Journal of Political Economy* 87(5), 972–990.
- Judge, G., Hill, R., Griffiths, W., Lütkepohl, H. and Lee, T. C. (1988): *Introduction to the Theory and Practice of Econometrics*, 2nd edition. New York, NY: Wiley.
- Kiker, B. F., Santos, M. C. and De Oliveira, M. M. (1997): Overeducation and under-education: Evidence for Portugal. *Economics of Education Review* 16(2), 111–125.
- Korpi, T. and Tåhlin, M. (2009): Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics*, 16(2), 183–193.

Manacorda, M. and Petrongolo, B. (1998): Skill mismatch and unemployment in OECD countries. *Economica* 66(262), 181–207.

McGoldrick, K. M. and Robst, J. (1996): Gender differences in overeducation: A test of the theory of differential overqualification. *The American Economic Review* 86(2), 280–284.

Nielsen, C. P. (2011): Immigrant over-education: Evidence from Denmark. *Journal of Population Economics* 24, 499–520.

Ortiz, L. and Kucel, A. (2008): Do fields of study matter for over-education? The cases of Spain and Germany. *International Journal of Comparative Sociology* 49(4–5), 305–327.

Ortiz, L. (2010): Not the right job, but a secure one: Over-education and temporary employment in France, Italy and Spain. *Work. Employment and Society*, 24(1), 47–64.

Ramos, R. and Sanromá, E. (2011): Overeducation and local labour markets in Spain. IZA Discussion paper series No. 6028, 1–17.

Sattinger, M. (1993): Assignment models of the distribution of earnings. *Journal of Economic Literature* 31(2), 831–880.

Shavit, Y. and Müller, W. (1998): *From School to Work. A Comparative Study of Educational Qualifications and Occupational Destinations*. Cary, NC: Oxford University Press.

Sianesi, B. and van Reenen, J. (2003): The returns to education: macroeconomics. *Journal of Economic Surveys* 17(2), 157–200.

Sicherman, N. and Galor, O. (1990): A theory of career mobility. *Journal of Political Economy* 98(1), 169–192.

Sicherman, N. (1991): “Overeducation” in the Labor Market. *Journal of Labor Economics* 9(2), 101–122.

Sloane, P. J., Battu, H. and Seaman, P. T. (1999): Overeducation, undereducation and the British labour market. *Applied Economics* 31(11), 1437–1453.

Tarvid, A. (2012): Effects of the field of higher education on the prevalence of over-education in European countries. *Education: Evaluation, Reform and Policy*, 167–184.

Thurow, L. C. (1975): *Generating Inequality: Mechanisms of Distribution in the U. S. Economy*. New York, NY: Macmillan Interactive Publishing.

Tsang, M. C., Levin, H. M. (1985): The economics of overeducation. *Economics of Education Review* 4(2), 93–104.

Tsang, M. C., Rumberger, R. W. and Levin, H. M. (1991): The impact of surplus schooling on worker productivity. *Industrial Relations: A Journal of Economy and Society* 30(2), 209–228.

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

Verhaest, D. and Omey, E. (2006): Discriminating between alternative measures of overeducation. *Applied Economics* 38(18), 2113–2120.

Verwiebe, R., Wiesböck, L. and Teitzer, R. (2014): New forms of intra-European migration, labour market dynamics and social inequality in Europe. *Migration Letters*, 11(2), 125–136.

Weeden, K. A. (2002): Why do some occupations pay more than others? Social closure and earnings inequality in the United States. *American Journal of Sociology*, 108(1), 55–101.

MODULE 2: GENDER PAY GAP – SES (2010)

1. Introduction

The persistence of gender differences in wages belongs to the best documented facts in labour economics. It has been motivation for a tremendous body of work analysing its roots and implications. Despite the variety of research approaches, many facets of the gender gap are still insufficiently explored. This is mainly due to the enormous behavioural complexity created by interlinkages between a person's work- and family-related decisions. Without a profound understanding of the causes of observed wage discrepancies, however, policymakers are unable to design the right policy mix for addressing the issue. The analysis of large amounts of individual data on job and worker characteristics is a necessary step on this way. This has led the European Union to promote research in this area. By exploiting data from the large-scale EU Structure of Earnings Survey (SES), a detailed picture of wage inequality in Europe can be obtained. According to the last estimations from 2013, the gender gap in hourly earnings is estimated to be 16.3% for the EU-28, with considerable differences at country level.⁶ To interpret these figures, researchers have developed decomposition techniques that attribute fractions of the gap to gender differences in certain observed characteristics. In this way, adjusted gaps are attained, representing the unexplained part of gender wage differences.

Our study contributes to this literature in several ways. First, we provide an update of existing figures on the unadjusted and adjusted gender pay gaps in EU countries based on the most recent wave of the SES. As a decomposition method, we apply the most well-known, Oaxaca-Blinder method (Oaxaca, 1973; Blinder, 1973). In this way, we are as close as possible to the methodology Eurostat employs when calculating and decomposing national wage gaps. Second, we enrich the literature by undertaking comprehensive country comparisons of the gap components. Overall, we analyse 21 EU countries (plus Norway), which clearly exceeds the scope of existing microdata studies (e.g. Arulampalam et al., 2007; Simón, 2012). Third, we differ from other studies in that we also examine and compare the sources of the unexplained gap, thus providing additional insights into the sources of the pay differential. Finally, we discuss our decomposition results in the broader context of female labour market participation, pointing to the role of selection effects and statistically unobserved gender segregation in industries and occupations.

Our findings confirm the persistence of gender wage discrepancies in Europe. The estimated unadjusted pay gap amounts to 15.3% in our cross-country analysis. In line with previous estimations, we detect considerable country heterogeneity. This heterogeneity is not limited to the size of the unadjusted gap, but also concerns its composition. The explained gap is estimated to be negative in six countries, while it

6 http://ec.europa.eu/justice/gender-equality/gender-pay-gap/situation-europe/index_en.htm

reaches levels up to 15% in other countries. In respect of the contributions of single characteristics, gender differences in the sorting into industries and into atypical employment (part-time work, temporary jobs) are predominantly widening the gap. Differences in educational levels and firm characteristics mitigate the gap. At the same time, the unexplained gap is nowhere found to be smaller than 5%, pointing to an important role of forces independent of observable worker and job characteristics. Moreover, our further analysis reveals a close positive relationship between the size of the pay gaps and female employment rates. This suggests the occurrence of positive selection into employment, a factor that limits the meaningfulness of observed pay gaps for an evaluation of female earnings perspectives.

The outline of the remainder of the study is as follows. The next section provides a brief overview of the theories and empirical results regarding size and components of gender pay gaps. Section 3 describes the measurement method and the data, and section 4 the model set-up. The results are discussed in section 5, and section 6 concludes.

2. Literature

2.1 Theoretical motivation of the gender wage divide

The literature on the determinants of gender gaps in average payment has produced an extensive set of theories helping to explain the persistence of the phenomenon. Comparing these contributions, perspectives differ substantially especially concerning one question: the relative role of individual versus societal explanatory factors. Human capital theory addresses explanations at the individual level, i.e. which are rooted in gender differences in work biographies. The impact of motherhood on labour supply lies at the centre of this reasoning. In a direct manner, researchers refer to the wage penalty working mothers receive due to a birth- and childcare-related absence from the labour market. This temporary absence can entail a devaluation of their human capital compared to men of similar age, especially with respect to experience-related knowledge (Becker, 1985). Hence, a pay gap attributable to this channel is simply seen as a reflection of observed productivity differences. Moreover, Becker (1985) and Fuchs (1989) speculate that most of the wage gap not attributable to experience is due to unmeasured differences between men and women in their commitment to parenting. One implication of a stronger commitment by mothers is a higher incidence of part-time jobs among female workers. For several reasons, part-time work is perceived as less efficient than full-time work from an employer's perspective. Dividing the same amount of work among more persons raises internal coordination costs. Moreover, training costs in relation to the actual workload are also higher. As a consequence, part-time workers are paid less than full-time workers with similar experience and education (Kalleberg, 2000). In the opinion of some Human Capital Theorists, another implication of gender roles in parenting could be a lower wage of mothers, as part of their energy has to be devoted to childcare. Finally, child-related effects also disseminate through potential repercussions on education decisions. Goldin and Polachek (1987) argue that the anticipation of future career interruptions lowers the returns and thus the incentives of young women to invest into education and job-related training. This, in turn, lowers their earnings perspectives even before starting a family.

A second strand of theories focuses on the role of social norms and perceptions as explanations for a persistent gender gap. The most visible indicator for the presence of such norms in the workplace is occupational segregation. It describes the observation that women tend to cluster in specific occupations, a phenomenon common to all industrialised countries. These occupations are for the most part characterised by a lower pay than typical jobs for male workers, thereby contributing to a gender differential in earnings (Marini, 1989). The causes are however highly disputed. It is argued that the observed selection into certain occupations stems from differences in worker preferences. According to this view, female workers tend to have particularly strong preferences for jobs characterised by pleasant working conditions and/or high degrees of time flexibility. These non-pecuniary rewards of working are considered in the wage negotiations, reducing the average payment in these occupations. The part of the wage gap resulting from job selection thus neither represents a sign of market failure nor discrimination.

However, this line of argumentation does not account for the existence of other forms of gender segregation, for instance concerning the access to higher positions within an occupation (Reskin and Roos, 2009). These rather point to the role of social norms and tradition as factors influencing job choices. Women might (consciously or unconsciously) be guided towards jobs with a traditionally female image. According to the theory of cultural devaluation, this assignment represents a trap for female workers. It states that jobs in which female workers start to reach a certain share in the workforce will experience devaluation in prestige over time, due to traditional prejudice associated with women's work (Reskin, 2001). This devaluation is marked by lower relative payment and lower opportunities of advancement. However, this theory alone does not explain the persistence of the gender wage gap, as women would have a strong incentive to switch to jobs with a male image. Instead, it requires an additional (or alternative) influence on the side of labour demand. One such influence could be the occurrence of 'taste-based discrimination', a concept originating from Becker (1957). Some employers might have personal preferences to hire male workers; some workers might prefer to collaborate with male colleagues. Again, social norms might lie at the origin of such preferences. The consequence is that women can only enter certain jobs and positions by accepting a lower payment than their male counterparts. As long as this 'preference for discrimination' is only detectable at the producer level, market competition should enforce a crowding-out of discriminating producers in the long-run: they are inferior in efficiency to non-discriminating producers that focus their hiring policies purely on worker productivity. Hence, this theory alone is also insufficient to explain the stable wage differential between men and women observed in modern market economies.

Another form of tradition-based explanation is the power of male-dominated career networks. The stability of such limited-access networks over time could be justified on a psychological basis by the concept of homophily: people are more likely to interact with people with similar attributes, as group homogeneity helps to minimise coordination costs (McPherson et al., 2001). Accordingly, male networkers have an incentive (to some degree unconscious) to prevent women from entering their circles. Obviously, this theory primarily addresses a gender gap at the top end of the wage distribution.

Finally, there is a third form of explanation lying somewhere in between the previous two. If information on the individual productivities of applicants is insufficient, it might be rational for employers to resort to information on the distribution of productivities within the societal group the applicant can be assigned to. The result is that individuals with abilities above the group-specific average are discriminated based on a statistical effect, as they are judged not by their own statistically unobserved high performance, but by the performance of the average group member. This phenomenon has been termed statistical discrimination (Arrow, 1973). If childcare induces the average woman to gather less work-related experience over her lifetime and therefore reach a lower level of productivity, statistical discrimination does deprive a female worker compared to a male worker with the same employment history. In contrast to taste-based discrimination, this behaviour is not punished by the market process as long as beliefs on gender differences are well rooted in data and continuously updated. This implies statistical discrimination to be more stable and long-lasting. However, perceptions of group-specific productivities could also be based on social prejudice. In this case, the outcome is identical to that of taste-based discrimination: women as a whole are deprived, but discriminating firms are expected to be outcompeted on the market.

While data-founded statistical discrimination might not be seen as a market failure in a static sense, it should be seen as one in a dynamic view. It lowers young women's returns on skill accumulation through schooling and work experience, keeping incentives for human capital investment low and thereby maintaining the skill lag of women in a vicious cycle. Lack of relevant information on the side of employers is thus another source of a persistent gender pay gap. In this view, educational and workplace policies promoting human capital investments of female students and workers would be considered an appropriate remedy.

2.2 Empirical evidence on wage gap determinants

In the recent past, highly comprehensive empirical literature on evolution and determinants of gender wage gaps has emerged. Depending on personal research interest and disciplinary background, the focus of works varies greatly. This is no surprise, given that questions of wage inequality lie at the crossroads of several schools and disciplines like economics, sociology and social psychology. Any attempt of summarising the literature has to start with classifying approaches according to method and/or content. In the following, we apply a broad classification of works into macro- and micro-centred approaches. Macro-centred approaches deal with the institutional determinants of gender pay gaps, while micro-centred approaches are interested in individual determinants (i.e. differences in worker characteristics).

Institutional determinants of the gender pay gap

The existence of considerable country differences in the size of wage gaps has aroused a lot of research interest. The fact that these differences are also observed among countries at similar stages of economic and social development has motivated researchers to investigate the role of labour market policies and institutions as a source of explanation. In a complex interplay with social norms and culture, policies could exert a discriminating impact on lifecycle decisions of men and women, affecting quantity and quality of labour supply in a dynamic manner. An important part of lifecycle decision-making is family planning. In this regard, childbearing still represents a risk for female workers, as it usually implies a temporary absence from the workplace that is associated with a delay in career progress or even a loss of task-specific human capital. Family policies have therefore gained particular attention in the gender gap literature.

One series of papers analyses the role of parental leave policies. These include all measures providing compensation payments to parents for taking a time off work to take care for (or in case of maternity leave: giving birth to) their children. As Blau and Kahn (2003) point out, the impact of such policies on the wage gap is a priori unclear. On the one hand, they allow pregnant women to keep the ties to their firms, thereby raising incentives for investments into firm-specific human capital. On the other hand, they could also contribute to an increase in the duration of career time-outs, preserving the gap in work experience. In their sample of 22 countries, Blau and Kahn (2003) identify the net effect to be insignificant. Waldfogel (1998) finds in a sample for the US and Great Britain that women who had leave coverage and returned to work after childbirth received a wage premium that offset the negative wage effects of children. Ruhm (1998) estimates in a comparison of 16 EU countries that parental leave of long duration harms women's relative wages. Childcare support is another form of family-related labour market policy, comprising direct

and indirect forms of subsidising external childcare. Its expected impact is a reduction of the gender pay gap, because it allows women to devote sufficient time to their jobs. In a comparison of OECD countries, Arulampalam et al. (2007) find that countries offering more generous childcare support tend to exhibit a smaller wage gap at the bottom end of the wage distribution and a wider gap at the top. This seems to suggest that primarily low-qualified women benefit from these measures.

Moreover, the role of wage setting institutions for the emergence of a gender differential has been subject to intense analysis. One research question is whether women are better off under centralised or decentralised (firm level) wage setting. In the developed Western economies, centralised wage setting usually takes the form of collective bargaining between trade unions and associations of employers. One common feature of such wage regimes is that wage dispersion is lower than under decentralised bargaining (Wallerstein, 1999). This is something that on average benefits female workers, due to their larger shares in the lower quantiles of the wage distribution (Blau and Kahn, 2003). At the same time, wage setting through decentralised negotiations is also characterised by a higher competitive pressure. This could entail a stronger impact of productivity differences, implying higher wage penalties on a lack of qualification and labour market experience. Especially the latter effect could work against women. On the other hand, a higher competitive pressure could also help to eradicate any statistically unobserved form of gender discrimination, as discrimination becomes too costly and endangers a firm's survival (Meng and Meurs, 2004). Daly et al. (2006) find in a cross-country comparison that Australia and the UK, as the countries with the most decentralised wage setting, had the smallest gender pay gaps in the 1990s. Apparently, it was primarily a shift away from seniority-based wages structure that has helped women to catch up. Antonczyk et al. (2010) find in an investigation for West Germany that the ongoing drop in collective bargaining has contributed to an increase in overall wage inequality, but hardly had an effect on the gender gap.

Another series of papers deals with the influence of minimum wage legislation. In general, the purpose of minimum wages is to reduce wage inequality by raising the earnings of the worst paid workers. Hence, concerning the gender gap, a similar argumentation can be made as in case of a centralised wage setting: relative wages of women should increase as they can benefit to a larger extent from wage increases at the bottom end of the distribution. However, additional indirect statistical effects could arise from the interplay with employment. Following the introduction of a minimum wage, a share of least-productive workers might lose their jobs. This effect could potentially further raise average wages of women. Shannon (1996) analyses the effect of minimum wages for Canada and finds that they have helped in closing the gender gap, however mostly due to their adverse employment effects. Boll et al. (2015) conducted a simulation study for Germany. They expect that the introduction of a general minimum wage in Germany will reduce the average gender wage gap by 2.5 percentage points when not controlling for employment effects. Moreover, they argue that under realistic assumptions employment effects should not significantly change the calculation.

Finally, technological progress influences gender differences in pay primarily through the phenomenon of skill-biased technological change. This describes the increase in the relative productivity of skilled labour in the course of computerisation and digitalisation of the workplace. It implies a growing supply of white collar jobs at the cost of blue collar jobs. On average, female workers tend to benefit from this development, as they less frequently perform physically demanding routine tasks and at the same time have caught up with men in the accumulation of skills related to

information technologies. Allen (2001) observed that female wages grew more than male wages in industries that intensively used high-technology capital in the US in the 1980s. Black and Spitz-Oener (2010) estimate that the decline in routine tasks explains a fraction of 50% of the closing of the gender wage gap in West Germany during the 1980s and 1990s. These task changes occurred most intensively in those occupations where the use of computers spread most rapidly.

Individual determinants of the gender pay gap

Differences in career biographies between male and female workers constitute another reason for a prevailing wage gap. On average, female workers could possess different market-relevant endowments than their male counterparts, potentially justifying at least some share of the observed gap. On the other hand, even with similar endowments, worker biographies could play a role as aspects such as education and experience might be evaluated differently in the case of male and female workers. Such evaluation differences are commonly interpreted as signs of gender discrimination. The share of the wage gap attributable to endowment differences is therefore subject to intense debate. Existing evidence strongly varies between countries, time periods and datasets. For instance, using French firm data, Barnet-Verzat and Wolff (2008) conclude that differences in the rewards men and women received for their labour market characteristics remain low; differences in observed characteristics are much higher. Boll and Leppin (2015) come to the same result for Germany. In contrast, Albrecht et al. (2004) find for the Netherlands that the main share of the gap is due to differences between men and women in returns to endowments rather than to differences in endowment sizes.

Among the single endowment effects, work experience is naturally of high concern, due to the potentially long career time-outs of pregnant women (see previous section). Most investigations accounting for the effect of experience yield the result that it makes up a dominating share of the explainable wage gap. In this point, studies apparently confirm the view of the human capital theory (cf. chapter 2.1). For instance, Blau and Kahn (1997) in their work with US panel data estimate full-time work experience to account for almost the complete explained gap. Waldfogel (1998), by adding controls for the returns to marital and parental status, yields lower but still impressive shares of 30% to 40% in a sample including the US and Great Britain. Boll and Leppin (2015) likewise detect a significant contribution for Germany. Furthermore, work experience could also matter with respect to its timing. Light and Ureta (1995) found in an analysis for the US that about 12 per cent of the overall wage gap could be attributed to gender differences in the accumulation of experience at the beginning of the career. Apparently, an out-of-work status is especially harmful during early stages of work life due to the path breaking effect for future advancement. Nevertheless, one practical point of concern in estimating an experience effect is its likely endogeneity (Weichselbaumer and Winter-Ebmer, 2007). The fact that female workers were less successful in gathering experience could itself have been the product of statistically unobserved ability differences or gender discrimination in job access. These repercussions imply a danger of overestimating a causal discriminatory impact of work experience.

A somewhat related aspect is job tenure. A positive relationship between tenure and wage level is a well-established result of the literature (Brown, 1989). It could both be traced back to productivity increases over time through on-the-job-training and to a signalling role of long tenure as an indicator of a high matching efficiency. Insofar as women face longer periods of labour market absence (especially related

to maternity and parental leave), average job tenure tends to be lower for female workers. Therefore, one could expect tenure to boost the overall gender gap. Results from Macpherson and Hirsch (1995) point in this direction; they identify average tenure to be lower in female dominated occupations. Moreover, job tenure could also affect the gender gap through differences in the returns to tenure between men and women. Munasinghe et al. (2008) find evidence for higher returns for males. They offer several explanations for this, starting from lower incentives of female workers to invest in job-specific skills to gender discrimination concerning promotion opportunities.

Similar evidence has been gathered for gender differences in the extent of work. As discussed above, the over-representation of female workers in part-time jobs is expected to worsen the earnings position of women. Indeed, Manning and Petrongolo (2008) document the discrepancy in hourly earnings of full-time and part-time working women in Great Britain. At the same time, the authors stress the endogeneity of part-time employment. Less ambitious women with lower incentives to invest into human capital might also prefer to work part-time. Moreover, the distribution of part-time work is also a reflection of occupational choices. Wood et al. (1993) find in sample of graduates from law schools that differences in work hours account for a third of the gap when job setting variables are excluded, while this share reduces to a quarter with job setting. Most recently, Goldin (2014) argues that gender differences in working time arrangements are the last crucial barrier faced by attempts to close the gender pay gap in the US. Finally, the role of part-time employment interferes with the role of the public sector discussed further below, given that public services are commonly characterised by a large share of women in part-time jobs (Gornick and Jacobs, 1998).

Against the background of theories of occupational segregation outlined above, another much investigated question is the contribution of occupational sorting. Women might cluster in less well-paid jobs, a fact whose persistence could stem from the importance of social norms and role models. Wood et al. (1993) find in their sample that job setting accounted for one third of the gender pay gap. Petersen and Morgan (1995) using cross-industry data of the US identify within-job wage differences to be less important than differences across occupations. Barón and Cobb-Clark (2010) even find for Australia that occupational segregation was an advantage to women except for those in high-paid private sector jobs. In contrast, Hinz and Gartner (2005) assign in their estimations for German workers a dominating role to within-job wage variation. In interpreting these country differences, at least two aspects have to be taken into consideration. First, Bettio (2002) stresses the importance of distinguishing between horizontal segregation, i.e. men and women sort into different (task-specific) occupations, and vertical segregation, i.e. men and women hold different hierarchical positions within an occupational group. She argues that a reallocation of jobs such that women imitate the male distribution of occupational positions (counterfactual situation) would notably reduce the pay gap whereas a replication of the male occupational distribution for females would have only a marginal effect for some countries and even increase the pay gap in others. In other words, the removal of horizontal segregation would not necessarily decrease the pay gap in Europe.

Hence, the interpretation of estimation results hinges on the question of whether or not the chosen classification of occupations also includes a vertical component. If it does, the impact of horizontal segregation is likely to be overestimated, as part of the observed variation stems from gender differences in hierarchical positioning. Second, estimates based purely on worker samples can be viewed as biased, as

they ignore interlinkages with the employment decision. Bettio (2002) sheds light on this relationship by drawing attention to the close correlation between the level of occupational segregation and female employment rates at country level in Europe. While Nordic countries are traditionally characterised by high female labour participation and high occupational segregation, the opposite holds for countries from Southern Europe. Bettio (2002) argues that this pattern is mainly rooted in country differences in the marketability of certain traditionally female dominated service tasks like nursing, cleaning and clothing. In Southern Europe, these tasks are still largely executed at household level, implying that the measured level of segregation on the market is significantly smaller than the real one (i.e. including non-market employment). As a consequence, the contribution of horizontal segregation to a gender gap in expected earnings tends to be underestimated in these countries. In comparison, the incidence of vertical segregation is more homogeneous. In a cross-national study, Wright et al. (1995) present evidence for significant hierarchical sorting in all countries examined. Busch and Holst (2009) find for Germany that gender-related selection effects have a crucial impact on the gender pay gap in managerial positions. In the same vein, in a cross-generational analysis Aisenbrey and Brückner (2008) detect that within-occupation stratification has, over time, become a more and more important factor in explaining the gender wage gap in Germany.

Furthermore, gender-related sorting is not necessarily limited to occupations, but can also occur across firms and industries. In this regard, some works have focused on differences between privately and publicly owned firms. Typically, the mean gender wage gap is considerably smaller in the public sector (Arulampalam et al., 2007). Gregory and Borland (1999) argue that these differences in wage structure are not surprising given that wage setting in the public sector occurs in a political environment, whereas private-sector decision-making occurs in a market environment. Moreover, anti-discrimination legislation may be more aggressively enforced in the public sector. At the same time, the public sector offers attractive employment conditions for mothers, due to the high degrees of protection, of time flexibility and of tolerance towards periods of absence (Kolberg, 1991). As a consequence, women tend to be over-represented in public sector jobs (Gornick and Jacobs, 1998). Nevertheless, there are also fears that this clustering of women could harm their economic opportunities in the long-run. The convenience of the working conditions might prevent them from accessing high-paid positions in the private sector, thereby stabilising vertical segregation (Mandel and Semyonov, 2005).

3. Measurement and Data

Over the years, researchers have developed several methods to analyse a gender gap in wages. Starting with the seminal work of Oaxaca (1973) and Blinder (1973), decomposition approaches have become the most popular tools. Their common idea is to split the observed gap into several parts, which are assigned a meaningful economic interpretation. In this way, the impact factors underlying the gap are distilled and assessed with respect to the magnitude of their contribution to the overall pay gap. The single approaches differ in two respects. First, the aggregate gap itself is defined in different ways. The classic Oaxaca-Blinder decomposition focuses on the gap in average hourly earnings between male and female workers. Other approaches undertake gender comparisons at different quantiles of the wage distribution (e.g. Albrecht et al., 2003) or make use of measures from the poverty literature (Del Río et al., 2011). Second, researchers favour different decomposition techniques. In the original Oaxaca-Blinder approach, a static decomposition of the current gap is in part explained by differences in worker characteristics and a remaining unexplained part is performed. Alternatively, Juhn et al. (1993) proposed to decompose changes in the wage gap over time into a portion due to gender-specific factors and a portion due to changes in the overall level of wage inequality. Moreover, several semiparametric techniques have been developed and implemented (DiNardo et al., 1996; Firpo et al., 2007).

For our estimations, we prefer to stick to the original Oaxaca-Blinder-model, both because of its widespread use in official statistics and its relative simplicity. Our strategy can therefore be summarised as follows: first, we compute the gender gap in average hourly wages for the aggregate sample as well as at country level. Then, an Oaxaca-Blinder-decomposition of these gaps into explained and unexplained parts is executed (for an introduction to this decomposition method, see next section or Methodological Guides). In this process, a series of worker characteristics included in our dataset is used as explanatory factors for gender differences in wage levels. Finally, the composition of the explained parts is analysed and compared across countries, i.e. the contribution of the single worker characteristics to the wage gap is discussed.

Our dataset consists of the most recent (2010) wave of the EU Structure of Earnings Survey (SES). The SES is a large enterprise sample survey providing detailed information on the relationships between the level of remuneration and individual characteristics of employees (sex, age, occupation, length of service, highest educational level attained, etc.) and those of their employer (economic activity, size and location of the enterprise). The national statistical institutes are responsible for selecting the sample, preparing the questionnaires, conducting the survey and forwarding the results to Eurostat. The sample regularly includes enterprises that have at least 10 employees and which are from sections C to O of the Statistical Classification of Economic Activities in the European Community (commonly referred to as NACE). However, public administration is excluded in some countries, which induces us to drop employees from this sector in our analysis. As further restrictions, no self-employed are included and information on sectors and occupational groups are only available at a limited level of disaggregation.

Given that data availability concerning individual and job-related characteristics differs to some extent between countries, we had to weigh the aim of accounting for as many insightful characteristics as possible against the need to preserve a sufficient number of countries for our analysis. In the end, we were left with 22 countries (21 EU countries plus Norway).⁷ The total number of observations is 8,829,191. In the following, the explanatory variables are described as individual worker characteristics, age and education were included. Age is measured in terms of six categories, where the youngest group comprises the 14–19 years old workers and the oldest group the more than 60 years old. The measure of education is derived from an aggregation of ISCED levels into three categories (ISCED 0–2, ISCED 3–4, ISCED 5–6). As job-related characteristics, contract type, firm tenure, hours of work, occupational group as well as branch, ownership and size of the enterprise were taken into account. Contract type is captured by a dummy variable that is equal to one for temporary and zero for permanent contracts. Firm tenure is split into four time spans (0–1 years, 2–4 years, 15–24 years, > 24 years). Hours of work are also only available as a categorical measure, distinguishing between full-time workers, those who work 60–99% and those who work less than 60% of a full-time worker's normal workload. Occupational groups are identified based on the ISCO-08 classification at the two-digit-level, discriminating between 42 different groups. The branch of the enterprise is assigned based on an aggregation of the NACE-Rev.2 classification, motivated by the need for cross-country harmonisation. It allows us to distinguish between 16 different sectors. For the impact of ownership, we include a dummy variable that is set equal to one if the firm is under public control. This is defined to be the case if a share of more than 50% is in public ownership. The size of the enterprise is measured by its number of employees, broadly categorised into enterprises with more and enterprises with fewer than 50 employees. Finally, in the estimations based on the aggregate cross-country sample, we add country fixed effects to the wage regressions. The country fixed effects capture statistically unobserved heterogeneity across countries.

7 Missing EU countries: Austria, Cyprus, Denmark, Ireland, Luxembourg, Malta, Slovenia.

4. Model

Formally, the Oaxaca-Blinder-decomposition consists of two estimation steps. As a first step, estimations of the determinants of hourly wages are carried out separately for male (m) and female (f) workers. This takes the form of separate wage regressions (for a brief note on the regression method see Methodological Guides). In a log-linear model, logarithmised hourly wages (W) are regressed on a set of explanatory factors, i.e. a range of worker and job-related characteristics (X). In the language of the literature, these characteristics are also termed *endowments*, as they are viewed as observable indicators of productivity differences partly explaining the wage gap. Formally, the regression equations look as follows (with β^j representing the estimated coefficient of the characteristic indexed with j and ε representing a residual term):

$$\ln W_{m;i} = \beta_m^0 + \sum_j \beta_m^j X_{m;i}^j + \varepsilon_{m;i}$$

$$\ln W_{f;i} = \beta_f^0 + \sum_j \beta_f^j X_{f;i}^j + \varepsilon_{f;i}$$

Afterwards, the resulting coefficient estimates are used to decompose the gender difference in the average wage levels (\bar{W}). This is achieved by replacing gender-specific log mean wages by the right-hand side of the two equations above. Following Blinder (1973), rearranging terms leads to the following expression:

$$\ln \bar{W}_{mm} - \ln \bar{W}_{ff} = \sum_j (\bar{X}_m^j - \bar{X}_f^j) \beta_m^j + \sum_j (\beta_{mf}^j - \beta_f^j) \bar{X}_f^j + (\beta_{mf}^0 - \beta_f^0)$$

The overall gender gap in log mean wages is thus split into three components. The first component represents a weighted sum of gender differences in observed characteristics, where each characteristic is weighted with the corresponding coefficient estimated for male workers. Economically, this term thus represents the part of the wage gap attributable to gender differences in observed endowments. It is therefore termed the *characteristics effect* (or *endowment effect*). For the analysis, the characteristics effect can be further decomposed into the contributions of the single characteristics. The second component is the weighted sum of gender differences in estimated coefficients, where the female endowments are the weighting factors. It shows which part of the wage gap is due to the fact that the same endowment generates different market returns for male and female workers. More precisely, it measures the change in the wage gap that would occur if the female endowment would be subject to the rewards estimated for men. Again, the contributions of single characteristics to this effect can be determined. Finally, the third component represents a constant term. It captures the influence of all statistically unobserved wage determinants on the gender wage gap, such as personal ability, negotiating skills and institutional setting. The sum of second and third component is termed *the coefficients effect*. It represents the unexplained part of the gender wage gap, as it cannot be traced back to observed endowment differences.

In the literature, the unexplained part is sometimes interpreted as an indicator of the extent of gender discrimination in payment (e.g. Del Río et al., 2011). This is however misleading in several respects. On the one hand, the fact that the unexplained part also comprises the influence of endowment differences in statistically unobserved characteristics between male and female workers could induce this indicator to overestimate the real level of discrimination. On the other hand, discrimination is likely to influence female workers' incentives to accumulate skill and therefore the *characteristics effect*. For instance, it could represent a contributing factor to gender differences in work experience. In this regard, the unexplained part will tend to underestimate the real extent of gender discrimination. Moreover, in interpreting the decomposition results, one has to be aware that endowment differences are evaluated with coefficients from the male wage regression and thus with labour market returns for male workers. The relative magnitudes of contributions to the *characteristics effect* therefore allow only limited conclusions on the effects of an elimination of endowment differences on the pay gap. This is of particular concern for characteristics whose wage effects are of opposite sign for male and female workers.

Furthermore, we need to stress that our approach does not include a selection correction when estimating individuals' earnings. Heckman (1979) and Lewbel (2007) established estimation procedures that take selection into labour market participation into account. We consider this issue relevant since women and men might be differently selected into employment, resulting in inconsistently estimated wages. However, as argued above, such an investigation fell apart from the intended update of the pay gap calculations according to the methodology used by Eurostat. So we end up with stating that our analysis focuses on wages of the employed only. As a consequence, no speculations on the counterfactual wage distributions for women and men in the absence of employment selection are made. This has to be left for further investigations.

5. Results

5.1 Decomposition in explained and unexplained gender pay gap

As a first result, we measure the cross-country gap in average wages of men and women to be about 15.3% (2010). This number is slightly lower than the 16.3% published by Eurostat for the EU-28 in 2013.⁸ Apart from the different time of measurement, this discrepancy is explicable by data constraints: in order to account for a sufficient number of explanatory factors in the decomposition, we were forced – among the other restrictions mentioned above – to limit the geographical coverage in our estimations to a subset of 22 European countries (including non-EU country Norway). As these figures do not account for gender differences in worker characteristics, they will subsequently be termed the unadjusted gaps. Applying the decomposition method outlined above, we find that about one third of the gap can be traced back to the role of the explanatory factors included in our analysis. A wage difference of 10.9% remains as the unexplained gap. Hence, the largest part of the gap is not due to gender differences in measured worker attributes.

However, at country level the picture varies drastically, as shown in Table 3. Looking at the unadjusted gap, figures range from 3.6% for Poland to 25.1% for Estonia. From a geographical perspective, it is noticeable that most Middle and Eastern European states are exhibiting gaps clearly below average, with the Czech Republic, Slovakia and Estonia marking the exceptions. Among the West European countries, only Italy is exhibiting a very small gap (4.5%). Further country variation is revealed by the decomposition results. The country ranking with respect to the unexplained gap changes substantially compared to the unadjusted gap. The role of gender differences in average worker features is in some countries not only more pronounced than in others, it also works in opposite directions. For instance, it is striking that the three countries with the smallest unadjusted gap (Poland, Italy, Croatia) all exhibit negative explained gaps (i.e. female workers would earn more than male workers according to the explained gap alone). Hence, the average female worker in these countries is endowed with better characteristics than her male counterpart, at least concerning those characteristics included in our dataset. The reason why also in Poland and Italy female workers nevertheless have lower average earnings is exclusively to be found in the unexplained residual.

⁸ <http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&plugin=1&pcode=tsdsc340&language=en>

Table 3: Unadjusted, explained and unexplained gender pay gap based on SES 2010 data, in %

Country	Unadjusted gap	Explained gap	Unexplained gap (adj.)	Country	Unadjusted gap	Explained gap	Unexplained gap (adj.)
Belgium	8.5	2.8	5.8	Latvia	7.4	-3.2	10.6
Bulgaria	8.6	1.3	7.3	Lithuania	5.8	-8.1	13.9
Croatia	5.7	-6.0	11.7	Netherlands	15.2	7.2	8.0
Czech Republic	16.5	3.4	13.1	Norway	14.3	7.6	6.7
Estonia	25.1	10.2	14.9	Poland	3.6	-7.8	11.4
Finland	20.7	9.4	11.4	Portugal	11.4	-0.9	12.3
France	13.5	4.8	8.7	Romania	7.1	0.8	6.2
Germany	22.2	14.5	7.7	Slovakia	16.6	2.2	14.4
Greece	13.1	5.5	7.6	Spain	17.4	5.4	12.0
Hungary	8.4	0.5	7.9	Sweden	14.0	6.3	7.7
Italy	4.5	-6.2	10.7	UK	20.3	6.0	14.3
				<i>Total</i>	<i>15.3</i>	<i>4.4</i>	<i>10.9</i>

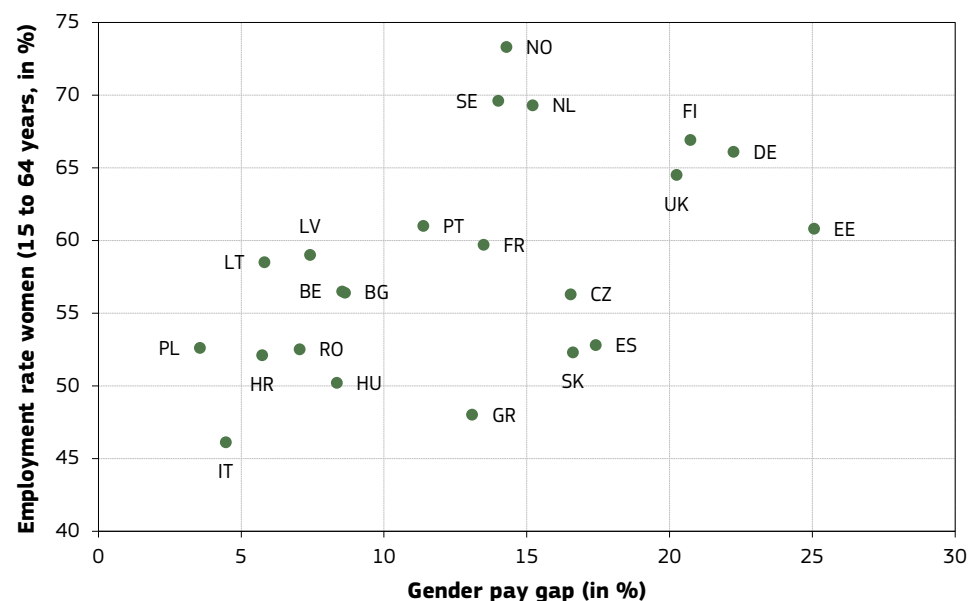
Sources: SES(2010), HWWI (2015).

Moreover, this unexplained part is nowhere identified to be negative. It does not even get lower than five per cent. It is thus this term that comprises the bulk of factors that prevent women from catching up. As explained above, it consists of two different kinds of effects. First, it acknowledges that the same endowment could be evaluated differently by the market, depending on whether the person is male or female. Second, it includes the impact of gender differences in those market-relevant characteristics not controlled for in our model. This second aspect is of special relevance, as our dataset does not allow us to assess potentially important gender differences related to actual work experience. It is interesting to see that some of the countries with negative explained gaps like Poland and Portugal perform worse than the country average when it comes to the unexplained gap. Apparently, one cannot conclude from that fact that women outperform men in attributes like education on a more egalitarian labour market environment in general. This provides justification for a more disaggregated analysis of the sources of the gender pay gap.

Before further decomposing the explained and unexplained gap into the contributions of single variables, it is important to shed light on a particular form of bias aggravating the interpretation. In analysing the gender gap, the comparison necessarily refers to actual wages paid to the working population, excluding the potential payment unemployed or inactive persons could receive. This creates a bias, given that the participation decision is also likely to depend on the potential earnings of the person. The expectation of low rewards could have motivated some share of women to stay outside the labour market, which would imply that we underestimate the real gender pay gap from the perspective of earnings potentials. Recently, in analysing US census data, Jacobsen et al. (2015) find evidence for a switch to such a positive selection during the last fifty years. If the opposite response occurs (as estimated by Beblo et al. (2003) for Germany), the implication is a corresponding overestimation. Moreover, such a bias does potentially not only concern the

cross-country gap, but also the comparison of country values. Figure 8 displays the unadjusted pay gaps together with the employment rates of women in the single countries. It documents a clear positive relationship between the two measures: countries with high female employment tend to exhibit high statistical pay gaps and vice versa. This seems to provide some confirmation for the positive selection hypothesis. Based on this, different clusters of countries can be identified.

Figure 8: Relationship between gender pay gap and female employment in SES



Sources: Eurostat (2015), SES (2010), HWWI (2015).

Taking a closer look, we can distinguish four country clusters, one group with low wage gaps and low female employment (henceforth named A), one group with medium wage gaps and low female employment (B), one group with medium wage gaps and high female employment (C) and finally one group with high wage gaps and high female employment (D). Beginning with A, this cluster is characterised by explained gaps below 3%, sometimes even below zero (HR, IT, LT, LV, PL, PT). That is, women's endowments rarely differ from men's or endowments are even in favour of women. The latter is the case in Italy and Poland. In both countries, occupational sorting is measured to be an important factor: women tend to cluster in better paid occupations. However, this results needs to be qualified against the background of employment selection and data limitations, as discussed below. Most countries that join cluster B are similar to A with respect to low endowment effects, but differ in the size of the unexplained gaps (CZ, GR, ES, FR, SK). Cluster C differs from B in that C countries exhibit high female employment rates. Countries of group C (NL, NO, SE) share an explained gap that amounts to roughly 7% and an equally high unexplained gap. This symmetry is special to group C. Finally, cluster D is found at the other end of the scale, with high female employment and high gender pay gaps (DE, EE, FI, UK). Here, Germany stands out with the highest explained gap (14%) in this

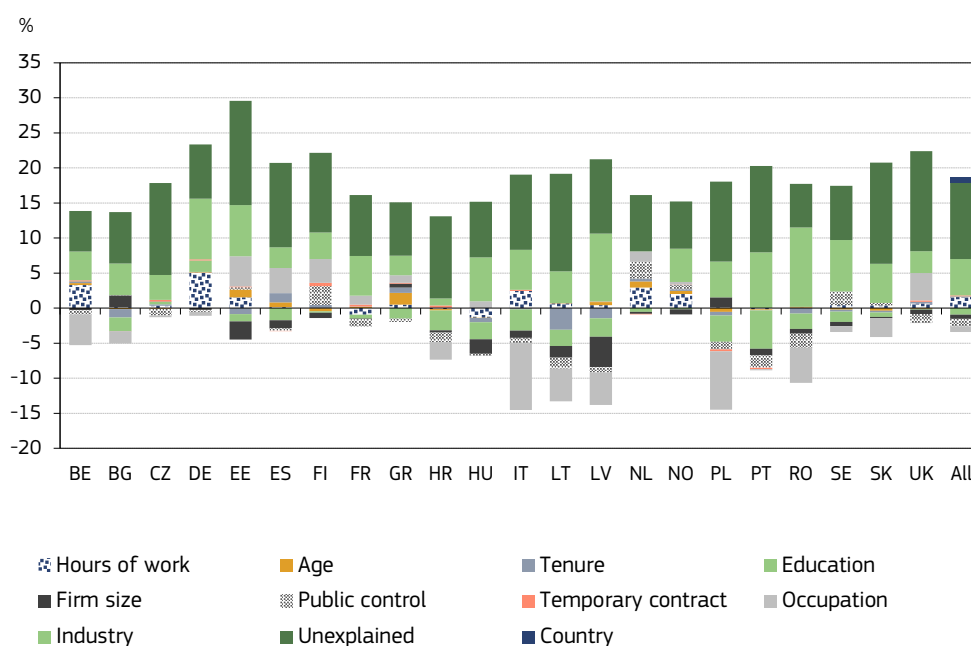
country comparison, being also the only country where the explained gap exceeds the unexplained gap.

All in all, countries with female employment rates lower than 60% have one thing in common, namely explained gaps below 5%. That is, a strong employment selection of women corresponds to low gender differences in pay-relevant endowments. Apparently, access to the market seems to substitute access to wage-relevant job attributes. In countries with low female labour market participation, only women with a high earnings potential access the market but after having passed this hurdle, these women apparently access attractive job attributes as easily as men. However, it has to be noted that they are worse paid for the same attributes and/or differ in unmeasured attributes from men. In countries with high access of women to the market (high female employment rates), women have more difficulties in accessing pay-attractive jobs than men, as reflected by endowment effects amounting to 6%–10%. This holds for Finland, the Netherlands, Norway, Sweden and the UK (Germany stands out with 14%). This is likely to be related to a high part-time share among employed women and/or more egalitarian gender roles. In five out of six of the aforementioned countries, at least 40% of all women aged 15–64 worked part-time (Eurostat-Data Explorer 2015; the exception refers to Finland where the respective quota amounted to 19% only). Thus, in all countries joining group *C* or *D* (except Estonia), egalitarian gender roles, mostly combined with family-compatible work arrangements enable women to enter the market, but this comes at the cost of a worse access to attractive job attributes (a higher endowment effect) compared to countries with female employment rates near 60% or below.

5.2 Decomposition of the explained gender pay gap

Figure 9 documents which share of the explained part of the gender pay gap can be attributed to which measured characteristic. Precise numbers can be found in Table A 13 in the Appendix.⁹ While some features show similar effects across countries, the role of others is highly heterogeneous.

⁹ Results of the wage regressions underlying our decompositions are available on request.

Figure 9: Decomposition of the gender pay gap (in %), 2010

Sources: SES (2010), HWWI (2015).

Something that can be noticed for all countries is that the selection of male and female workers into different sectors contributes to the existence of wage differences. Hence, a significant part of the gender gap is due to the fact that women are over-represented in industries with low pay levels (and accordingly under-represented in well-paid industries). This is consistent with recent results by Simón (2012) for the 2002 wave of the same dataset. In the cross-country sample, women are particularly over-represented in education as well as in health and social work activities. At the same time, they are highly under-represented in construction and in manufacturing sectors such as chemical products, electric and transport equipment.

In a country comparison, the largest effects of sectoral distribution are measured for Romania and Latvia, where its contribution to the overall gender gap amounts to 11.3% and 9.7%, respectively. In both countries, the comparatively small presence of women in well-paid jobs in the area of manufacturing and construction is again responsible for this result. At the other extreme, there are two countries where the industry effect remains fairly marginal: the Netherlands (<0.01%) and Croatia (0.01%). In the Netherlands, manufacturing sectors as well as wholesale trade are an important part of the explanation. Dutch women show a lower participation in these sectors than in cross-country average. At the same time, these sectors offered, all else being equal, a comparatively low remuneration compared to other sectors in the Dutch economy, a fact that primarily concerned men. In Croatia, the conditional wage bonus in construction is measurably lower than in the cross-country average. In general, we need to remain cautious with our interpretation, given that the real industry effect could be diluted by the congruence between occupational choice and sector.

Among the remaining characteristics effects, there is none that works in the same direction in each country. One that is at least almost homogeneous is the effect of

firm size. The fact that the genders are differently distributed across firms with different size mitigates the wage gap by 0.6% in the cross-country estimation. In the cross-country average, women work more often in large firms. Firms with more than 50 employees exhibit a higher share of female workers than smaller firms in the aggregate sample. In addition, the payment level in large firms is, *ceteris paribus*, higher – a result that is well documented in the labour economics literature (Oi and Idson, 1999). Explanations could be the occurrence of productivity gains through higher division of labour or the need to pay compensating differentials due to the unpleasantness of working in an impersonal atmosphere (Masters, 1969). As a consequence, the characteristics effect of firm size reduces the gender pay gap. The only conflicting evidence at country level is obtained for Bulgaria, Poland and Greece. In these countries, this is not caused by a reversal in the effect of firm size, but by the fact that women are under-represented in large firms. Nevertheless, the result for firm size might be sensitive to sample selection, given that very small firms with less than 10 employees are under-represented in the SES. We will reinvestigate the effect based on EU-SILC data in module 3.

Moreover, the role of schooling tends to contribute to wage convergence. It is an unsurprisingly unanimous result in our estimations that graduates from tertiary education received higher average earnings than workers at lower educational levels. A more insightful outcome is that female workers in most countries exhibited a higher average level of education than their male counterparts, at least when measured on our three-level scale. The consequence is a diminution of the cross-country gender gap by 0.9%, clearly exceeding previous results by Simón (2012). In two countries, Poland and Portugal, the diminution even exceeds 3%, mainly due to large gender differences in the shares of college graduates. On the other hand, with Germany we witness a case where differences in schooling boosted the wage gap further, by 1.6%. This is explained by the fact that unlike in most other countries German women still lagged behind regarding participation in tertiary education.

The form of economic control over the firm is another factor which predominantly reduces the gender wage differential. The fact that male and female workers are unequally distributed between private and public companies helps to narrow the gap. In all observed countries at the given point in time, female workers were over-represented in publicly controlled firms. Given that we controlled for sector and occupation, this is not simply to be seen as a reflection of occupational preferences. At the same time, conditional remuneration was, in the majority of countries, higher in public than in private firms, implying a reduction of the wage gap by 1.1% in the aggregate and up to 2% (Romania) at country level. Among the opposite outliers, Finland and the Netherlands stand out. Here, working in the public sector implied a wage penalty, yielding an increase in the gender gap by 2.7% and 2.4%, respectively.

In contrast, a job characteristic that predominantly raises the wage gap is hours of work. In all countries under observation, female workers have more often been employed part-time than male workers. In most of them, part-time work was, *ceteris paribus*, associated with lower hourly earnings. As discussed in the literature section, this can be related to perceived coordination costs and restrictions in the access to internal training. In all, it contributes to a widening of the cross-country gender pay gap by 1.6%. This fits recent evidence by Goldin (2014) for the US, who assigns working time arrangements a key role for explaining the incomplete gender convergence on the US labour market. An outlier concerning the magnitude of this effect is Germany, where it reaches a level of 5.0%, the second largest of all measured characteristics effects in this country. This is mainly attributable to the

pronounced gender differences in part-time work of low scale (less than 60% of a full-timer's normal hours). Exceptions in the other direction are France and Hungary, where the effect is slightly negative, due to the surprising result that part-time workers experienced a wage bonus in these countries.

Another channel that tends to widen the gender gap is the distribution of temporary vs. permanent contracts. Working in a temporary position reduces the expected earnings in almost all country regressions. This is consistent with general findings of the literature (Booth et al., 2002). Temporary workers have less incentive to accumulate job-specific human capital, as they face the risk of depreciation when the contract is not prolonged. For the same reason, employers are also less inclined to give them access to internal training. Temporary workers are also likely to represent a selection in the sense that they tend to be unsure about their future career and are therefore generally less willing to make any specific human capital investments. The outcome is a lower payment due to lack of specific skills. In turn, this contributes to the wage gap because temporary positions are more frequent among female workers in the majority of countries. This seems intuitive in the presence of self-selection: facing a higher risk of career interruptions through childbirth, women on average are less inclined to commit to a certain career path. Nevertheless, the overall effect remains of low magnitude. In our cross-country sample, temporary work widens the wage gap by only 0.1%. At country level, the maximum contribution is 0.5% (Finland). Cases where the effect goes in the other direction comprise those countries where the gender distribution of temporary work is reversed. In Poland and Portugal, this results in a modest reduction in the gender pay gap by 0.3% and 0.2%, respectively.

The role of the remaining characteristics is highly ambiguous in country comparison. First, this concerns workers' age distribution. To ensure a correct interpretation, we need to emphasise here that the measured impact of age is just like the impacts of other characteristics – a *ceteris paribus* effect, i.e. it is conditional on all else being equal. That is, the single effect of age arising from our estimations measures the wage difference between women and men of different age who exhibit the same qualification, experience, occupation and other wage-relevant characteristics. Therefore, this effect must not be confused with age-specific wage profiles for men and women that usually feature an increasing wage gap with increasing age. This is mainly due to the fact that in these cross-sectional illustrations, age simply carries (not displayed) other relevant factors like experience that notably differs between men and women. In the aggregate estimation, the net effect of age differences is practically zero (0.02%). Effects of the single age groups are of a similar magnitude. A look at the wage regressions shows this is not due to an irrelevance of the factor age in wage setting. Compared to the reference group of 40–49 year old workers, workers in most other age groups are estimated to receive significantly lower earnings in the cross-country regression for male workers, reproducing the typical inversely U-shaped wage evolution from the literature (Skirbekk, 2004). Rather, differences in the age distribution of male and female workers are simply too small to let this affect the wage gap. Nevertheless, this cross-country average does not adequately describe the situation in many single countries. On the one hand, we see a country like Greece where gender differences in the age distribution of workers are estimated to raise the gender pay gap by 1.7%. This results from the interplay of two factors. Greek female workers are on average younger than their male counterparts. At the same time, seniority is apparently a more important factor for wage setting in Greece than in other EU countries. The oldest group of male workers (60+) is here measured to receive a wage bonus compared to the reference

of 40–49 years old. On the other hand, we have a country like Poland, where age differences reduce the gap by 0.5%. Here, we observe an inversely U-shaped wage structure. Moreover, female workers in Poland show higher relative frequencies than men in the best-paid group of 40–49 years old.

A second highly ambiguous effect is measured for firm tenure. In the aggregate sample, differences in tenure raise the gender pay gap by a mere 0.1%, which is significantly lower than the 0.5% estimated by Simón (2012) for his dataset of nine European countries. In line with basic intuition, longer job tenure is associated with higher earnings in the cross-country regression. This can both be explained by a mechanism of self-selection (higher wages imply higher job satisfaction, thus workers stay longer) and the productivity-enhancing accumulation of job-specific human capital over time (Topel, 1991). Moreover, women exhibit a slightly larger average tenure in the cross-country sample. At country level, the positive relationship between wages and tenure is generally confirmed. Differences arise from the correlation of tenure and gender. In Finland, firm tenure contributes to the overall wage gap with 0.5%. Here, this can be traced back to a comparatively high share of female workers with short tenure. In Bulgaria, on the other hand, the impact of tenure on the wage gap is measured as - 1.2%, reflecting a high local share of female workers with very long tenure.

Finally, the characteristic causing the most heterogeneous effects is occupation. Its contribution to the gender pay gap in the aggregate sample is -0.8%. Hence, at the time of observation, women tended to cluster in the better paid occupational groups (from a male perspective). At a first sight, this seems to reject the theories linking occupational segregation to gender pay differences laid out in the previous section. However, we need to remain cautious with our interpretation, due to several data limitations. First, we merely distinguish between 43 occupational groups, thereby not capturing the full extent of gender heterogeneity in occupational sorting. Second, we can expect a high degree of correlation between occupational choice and sector, up to the point that some occupations are only observed within some sectors. For the industry effect, we had seen that selection of male and female workers into different sectors almost uniformly raised the gender pay gap. Hence, some part of the effect of occupational segregation could have been attributed to industry differences. Thirdly, with the occupational classification at hand, it is not possible to adequately control for vertical hierarchy. This is an important point since the different allocation of women and men to hierarchical positions within occupations is a robust finding in the literature (e.g. Bettio and Verashchagina, 2009). Note that, according to the formula of the wage gap decomposition in chapter 4, the characteristics effect is the gender difference in occupational characteristics weighted with male returns to that specific characteristic. That is, a negative characteristics effect with respect to occupation signals that women are more frequently allocated to occupations with above-average (male) wages. However, the characteristics effect does not tell us anything about the wage differential between women and men in a particular occupation. Instead, the effect of vertical segregation on gendered pay is captured by the coefficient effect of occupation (see chapter 5.3 below). Last but not least, employment selection matters: In some countries, tasks associated with a female image are still largely executed outside the formal labour market (Bettio, 2002).

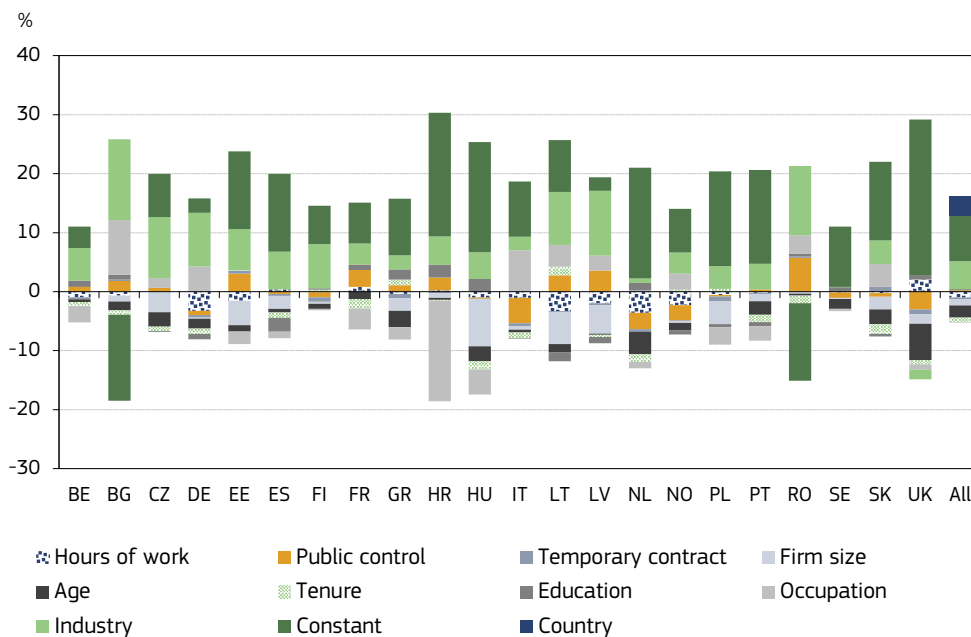
Referring to these particularities, the moderate effect measured for occupational endowment achieved from the aggregate sample appears a bit less striking, especially since it does not stand out in the literature (cf. Bettio and Verashchagina, 2009; Ministère du Travail, de l'Emploi, de la Formation Professionnelle et du Dialogue

Social 2015 for France). Moreover, the overall effect hides tremendous heterogeneity across countries. In Spain and the UK, occupational differences are measured to contribute more than 3.5% to the overall wage gap, implying this to be the prime factor responsible for the existence of a positive explained gap in these countries. In Italy and Poland, we witness a massive negative impact reaching levels of -9.5% and -8.3%, respectively, supporting the result that endowment differences in total work in favour of women. A detailed analysis of the contributions of the single occupational groups reveals this is mainly the outcome of a strong concentration of female workers in one group: teaching professionals. Again, this has to be interpreted in the context of employment selection. As displayed in Figure 8, Poland and Italy belong to the group of European countries with low wage gaps and comparatively low female employment rates. Apparently, this is a reflection of the fact that some typically low-paid service tasks like nursing and cleaning, which have traditionally been viewed as women's work, are to a large part still not delegated through formal work contracts in these countries, but mostly executed within households. Hence, they are not observed in the data. The result is a statistical reduction in the observed gender pay gap, which is ascribed to the occupation effect.

5.3 Factors behind the unexplained gender pay gap

Results in Table 3 have shown that the unexplained gap is positive everywhere and makes up the largest part of the overall gender wage gap in almost all countries under observation (with Germany and Norway marking the exceptions). Given the unavoidable data limitations, this does not come as a surprise. First and foremost, this results from the lack of a measure for actual work experience. An approximation by potential experience as measured by a worker's age has to remain highly imperfect, as it does not account for gender differences in labour market absence, especially related to birth and child caring. Endowment effects resulting from these differences are implicitly included in the residual gap. Moreover, it is also likely to include those effects of hierarchical and occupational sorting, which cannot be captured by the precision and aggregation level of an occupation measure like ours. Similar statistically unobserved effects could stem from factors like personal abilities and negotiating skills. For this reason, it is important to repeat our warning from above that the unexplained variation should not be interpreted as the extent of gender discrimination in wage setting.

Nevertheless, this does not mean that the sources of the unexplained gap have to remain completely in the dark. As a result of our decomposition method, the residual gap also includes the effect of a different evaluation (coefficient effect) of measured characteristics in the male and female subsamples (see Methodological Guidelines for further explanations). These effects can again be assessed in their magnitude for the single characteristics. Figure 10 plots the contributions to the unexplained gap at country level. Positive contributions imply that male workers receive a higher monetary reward for exhibiting the particular wage-relevant characteristic than female workers, a fact that widens the gender pay gap. Accordingly, negative contributions express the opposite. They imply that women are better paid for a particular characteristic than men, which mitigates the gender pay gap. Precise numbers can be found in Table A 15 in the Appendix.

Figure 10: Decomposition of the unexplained part of the gender pay gap (in %), 2010

Sources: SES (2010), HWWI (2015).

As it is the case for the characteristics effect, sources of the coefficients effect differ substantially between countries. Nevertheless, some major patterns can be identified. First, industry is estimated to exert a sizeable positive coefficients effect in almost all countries except the Netherlands, Sweden and the UK. For the aggregate sample, this effect equals 4.7%. In Belgium, Sweden and Romania, the magnitude even exceeds 10%. That is, there is a relative within-sector male wage premium in the female dominated sectors in all countries except the three named above. This indeed hints at considerable intra-sectoral gender heterogeneity with respect to the sorting into occupations and hierarchical positions. Apparently, much of the sorting takes place within rather than between industries. An implication is that the wage level effects of different wage setting regimes at industry level (e.g. centralised vs. decentralised, influence of trade unions) in most countries do not fully make up the pay effect of sorting although admittedly, such a conclusion has to be drawn with caution, due to the limited disaggregation of our industry variable. However, Goldin (2014) brings some light on the question of what the true sources of the unexplained gap might be about. In a nutshell, she argues that they are not driven by different human capital of men and women but by firm level differences in the cost of time flexibility. Hours worked are of different worth in different sectors and occupations, hence ignoring the interaction of sector with hours worked leaves a crucial source of pay differentials out of the analysis (leaving it to the ‘unexplained’ part or more concrete: the coefficient effect of industry). Based on American Community Survey data, Goldin shows that occupations of different sectors differ in their ability and cost to provide employees with reduced working hours on the occasion of family events. Technically, this results in a nonlinear relationship between earnings and hours of work: since deviations from the ‘full-time full-year’ (FTFY)-scheme are perceived as costly by the firm, long hours and continuous careers are honoured

with high wage premiums and short hours and/or career disruptions are strongly penalised. Due to a lack of compatibility with family tasks, women in particular give up employment at the occurrence of births and high wage penalties upon re-entry combined with further penalties for short hours result in high gender pay gaps in those industries. By contrast, industries which had successfully adapted to flexibility demands of their (not only female) workforce are shaped by almost linear earnings-hours worked relationships and hence smaller gender pay gaps. Goldin concludes that to further reduce the gap, sectors should strive to develop strategies to decrease the cost of time flexibility. We support this view in stating that policies focusing on the sector level aiming at closing the within-sector pay gap could be viewed as an important supplement to a cross-sectoral harmonisation of wage setting institutions in tackling the gender gap.

The second consistent pattern is the negative coefficients effect of the age composition. It reduces the gender pay gap by 2.1% in the aggregate sample. It is also negative throughout the single-country estimations, but not always significant. It reaches by far the strongest impact in the UK with -6.2%. The interpretation of these results strongly hinges upon the estimated distribution of wages over the lifecycle. Our cross-country wage regressions provide evidence for an inverse U-shaped relationship between age and wage level for male workers, with conditional pay being highest in the reference group of 40–49 years old. This not only finds supports in large parts of the labour economics literature (Skirbekk, 2004), it is also consistent with our result of a U-shaped relationship between age and overeducation risk in module 1. For female workers, instead, conditional wages are estimated to be continuously higher for older age groups in the cross-country estimation as well as in many single countries. Not being in the reference group of the middle-aged is, for young women, associated with a lower wage penalty than for young men and for older women even with a wage bonus. Thus, being evaluated like male workers would result here in an average wage loss for women, explaining the negative coefficients effect. Two reasons for these gender-specific age patterns seem plausible. First, statistically unobserved occupational sorting could again impair our estimations in the sense that women are under-represented in some physically highly demanding jobs in crafts and manufacturing. In executing these jobs, physical degradation in the process of ageing is coupled with a strong productivity decline. If this primarily harms earnings perspectives of old male workers, another channel through which occupational segregation affects the wage gap is established. Second, selection effects associated with the employment decision are another candidate. In module 1, we attributed the decline in the overeducation risk of medium-skilled women with increasing age to this channel. Here, the corresponding argumentation can be repeated. If with increasing age more and more women exit the labour market (i.e. become inactive) and the remaining active women represent on average a positive selection in terms of productivity-relevant characteristics, the statistical result of a beneficial age effect emerges. Even though we are not able to quantify the relevance of these selection processes, their pure existence requires that any policy recipes should be checked for their indirect impacts on women's employment decisions.

Finally, the constant term represents a major contributing factor in the majority of countries. With a contribution of 7.7% in the cross-country sample, it is almost exclusively responsible for the existence of an unexplained wage variation. It captures the influence of statistically unobserved variables in the data set at hand. As discussed above, gender differences in actual work experience over the lifecycle are expected to make up the bulk of this amount. The wage-reducing effect of a temporary labour market absence of women due to birth and childcare is nowhere explicit-

ly accounted for in our approach. However, this factor alone does not elucidate the occurrence of two striking exceptions at country level in this regard. In Bulgaria and Romania, the constant term is negative and smaller than –10%. Given the specific political and economic situation of these countries as still quite recently acceded EU member states, this asks for a detailed institutional analysis. Again, this points to the natural limits large-scale cross-country studies like ours face in disentangling the socioeconomic interrelations behind a certain observation. Illuminating these particularities requires further detailed country studies incorporating the interplay of individual decision-making and institutional backgrounds.

6. Conclusion

This study has investigated size and sources of gender wage gaps in the most recent wave of the EU-SES. Our first result was already a crucial one: a significant wage gap between male and female workers is still an undeniable reality in every single EU country under observation. Nevertheless, our wage decomposition analyses revealed a tremendous degree of country heterogeneity concerning the roots of this phenomenon. This holds in particular for the size of the gap that is attributable to gender differences in the measured wage-related worker and job characteristics. While this explained gap operates in some countries like Germany and Estonia decisively in favour of men, in others like Poland and Italy it is advantageous to women. An analysis of the relationship of the gender pay gap with female employment rates hints at selection effects as a major source of these country differences. Countries with low wage gaps tend to be characterised by low female labour market participation, suggesting the existence of a positive selection of working women with respect to productivity-relevant characteristics and/or a trade-off between family-compatibility and attractive pay. Therefore, one of our major conclusions is that in discussing wage gaps it has to be kept in mind that a more or less significant part of the female population is not in our sights. By utilising the correlation with employment rates, we alternatively form and analyse four clusters of countries differing in labour market perspectives for the female workforce.

From an examination of the contributions of the single observed characteristics, gender differences in the sorting into industries are identified as the strongest contributing factor by our decomposition method. This result is consistent with a previous study by Simón (2012). However, we need to be cautious with interpreting this effect, given its close relation with occupational choices and the limited level of disaggregation in our data. In many countries, the over-representation of women in atypical employment in the form of temporary and part-time jobs entails an additional wage penalty for women. On the other hand, factors that mitigate the pay gap in the majority of countries are the distribution of male and female workers into firms of different size as well as gender differences in schooling. The role of occupational segregation is measured to be highly heterogeneous across countries. This is explainable both by the biasing effect of its correlation with the industry variable, its interplay with female employment rates and the lack of control for vertical hierarchy. Finally, our results for the composition of the unexplained gap confirm our intuitions on the role of intra-sectoral pay equity and the role of selection effects. First, as Goldin (2014) points out, it is likely that sectors which do not manage to reduce the cost of time flexibility in terms of working hours and temporary employment breaks compensate their employees who stick to the full-time, full-year (FTFY) standard with high wage premiums, explaining the within-sector pay inequity with firms' personnel management rationales and not with human capital related differences between men and women in statistically observable 'endowments'. Goldin concludes – and we would adopt this suggestion based on our findings – that policies aiming at tackling the gender gap in pay should address both across-sector and within-sector differences in gendered pay. In the latter aspect, it seems particularly important to support sectors to develop strategies to decrease the cost of

time flexibility, instead of only striving to even the distribution of men and women over sectors and occupations. In this context, secondly, it will be important keep the selection issue on the screen. In a cross-country but also importantly in a lifetime perspective, the goal must be to eliminate the trade-off between low pay gaps and high female employment rates and instead combine attractive pay with time-flexible work arrangements.

Our results provide motivation for further investigations. To underpin the relevance of employment selection, estimations could be undertaken based on mixed sample of employed and unemployed/inactive persons. This would allow a modelling of employment decisions based on observed characteristics. By integrating this decision variable as an explanatory factor into the decomposition, a selection bias could be eliminated (Heckman, 1979). Moreover, an increase in explanatory power could be created by including additional characteristics in the decomposition, which was also impossible with the given dataset. This primarily concerns the lack of a measure of actual work experience. Other missing characteristics of interest refer to the household level. For instance, it would be interesting to know how far the income and educational status of a partner affect the wages of male and female workers differently. Lastly, in light of our analysis of overeducation in module 1, the implication of this phenomenon for the gender pay gap deserves attention. In module 3, we will make an attempt precisely in this direction. We will conduct the same decomposition analysis based on data from the European Union Statistics on Income and Living Conditions (EU-SILC), which allows us to exploit the additional information this dataset contains.

MODULE 3: GENDER PAY GAP – EU-SILC (2012)

1. Introduction

The preceding analysis in module 2 has demonstrated that the gender pay gap currently observed in Europe is the outcome of a range of intertwined factors. At the same time, a sizeable portion of the gap could not be explained at all with the worker characteristics used in the decomposition. Given the limited scope of the dataset in this respect, this does not come as a surprise. Mainly, information on a worker's household context, e.g. status and employment characteristics of a living partner, is missing. Existing research shows that a spouse's earnings and employment status is to some degree correlated with female labour market success (Henz and Sundström, 2001; Brynin and Francesconi, 2004; Shafer, 2011, Goldin, 2014). Note that as our target variable measures gross earnings, we focus on market income and abstain from fiscal effects arising from the tax and transfer system. Furthermore, we have to remind the interested reader that earnings are observed for the employed only, that is, a well-established body of literature that deals with the role of the household context in shaping female employment propensity is masked out here as well.

Thus, what are the theoretical underpinnings for the above cited empirical linkage between a woman's earnings and the composition of the household she lives in? First, also in this module we have to differentiate between household *endowments* that hinder or boost one's earnings perspectives, and evaluation effects in the sense that one characteristic pays off more for one person than others. Moreover, if one thinks of the working hours of women as a continuous spectrum including the null, the distinction between insiders and outsiders of the market vanishes and the focus is on employment intensity only. Assume that this intensity (that is reflected by further job-related characteristics like hierarchical position, occupational prestige etc.) is shaped by a woman's eagerness and ability to fully tap her earnings potential. Then, the woman's earnings appear as a measure of this multifacet 'intensity', and we may ask what household factors could motivate and hinder women to fully explore their inherent capacity.

In the analysis of realised earnings, economic incentives and institutional barriers play a crucial role. If one imposes the assumption of the unitary model (Samuelson, 1956; Becker, 1965) that partners pursue joint interests, the criterion of comparative advantage shapes the couple's labour division (Becker, 1965, 1981). For example, if the male spouse is relatively more productive in the market due to a higher qualification level and the female partner is relatively more productive in household tasks, a part-time job for the female partner may be the likely outcome. Bargaining theories (Manser and Brown, 1980; McElroy and Horney, 1981) that allow for diverging interests of partners rely on their human capital related outside options. In the previous example, the same result may arise from a different motivation:

The less educated partner is less powerful in intra-household bargaining processes and this is why she fails to advance her personal career. Here, the partners' formal education ratio serves as a 'sharing rule' (Browning and Chiappori, 1998) governing the division of partners' common resources. Looking at gender theories (Bittman et al. 2003; Brines 1994) provide a third explanation referring to gender stereotypes. Women might deliberately leave part of their economic resources untapped if this accords with their individual orientation in terms of gender roles. To turn the given example to the woman's advantage, assume that she holds a higher qualification than her husband. Then, she might be reluctant to put herself into the main earner position if this contrasts with prevailing gender roles in her social networks, particularly those held by her partner. Even in 2014, a special Eurobarometer revealed that most Europeans think that family life suffers when a mother has a full-time job and around half believe that men are less competent than women at performing household tasks (European Union, 2015). Finally, other institutional barriers like a lack of childcare amenities might hinder women with small children to make much money and to approach men's earnings. Or, putting the argument to the contrary, a lack of childcare facilities might pose prohibitive entry barriers on all but those women with very high earnings capacity, and cause in this sense a positively selected female employee group.

To sum up, different theoretical strands motivate a linkage between household context and individual earnings. In this regard, the European Statistics on Income and Living Conditions (EU-SILC) represents a promising alternative. Besides its fairly comprehensive coverage of partner-related features, it also allows us to extend our analysis to a larger set of countries.

The specific objective of this module 3 is therefore twofold. First, it aims to reproduce the decomposition analysis of the sources of the gender pay gap in Europe based on an alternative dataset, thereby checking the robustness of our previous results with respect to data coverage and sample selection. Second, we widen the perspective by considering additional explanatory factors for gender wage differences within our decomposition analysis. This provides us with a clearer picture of the forces at work. As part of this model extension, we also examine the effect of overeducation as a potential determinant. For its interpretation, we can draw upon our results from module 1, especially concerning gender differences in the likelihood of overeducation. In a sense, this closes the circle of our analysis. With regard to the interpretation of results of module 3, we need to stress that comparability to those of module 2 is limited. In addition to the general effect of using another dataset, the different selection of countries and the deviating set of explanatory factors marks a further departure. Moreover, the definition of wages as the target variable differs slightly between the datasets (see Measurement and Data).

The structure of this module is the following: Section 2 explains our sample composition and measurement techniques, stressing the differences to the previous module. Section 3 represents the results section. It first provides and discusses descriptive results for the gender pay gap based on the new dataset, then examines the role of the single determinants in the decomposition analysis. Section 4 concludes with summarising the main insights from module 3.

2. Measurement and Data

In this third module, we use data from the wave 2013 (rev. 1) of the European Union Statistics on Income and Living Conditions (EU-SILC). The objective of the EU-SILC is to provide data on income, poverty, social exclusion and living conditions in the European Union for persons aged 16 and above. The EU-SILC data is available as an (unbalanced) panel or as cross-sectional data. We use the latter version.

Our analysis is based on the anonymised microdata from the EU-SILC scientific use files. We consider the following 31 countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Iceland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland and the United Kingdom. We exclude Malta from the sample due to a non-comparable ISCO coding of occupations. Hence, in comparison to module 2, the countries added are Austria, Cyprus, Denmark, Ireland, Iceland, Luxembourg, Serbia, Slovenia and Switzerland. In preparation for our analysis, we imposed further restrictions on the data. In consistency with our previous gender pay gap calculation, we exclude apprentices and persons working in NACE sector O (Public administration and defence; compulsory social security). Apprentices are defined as persons attending ISCED levels 0–4.¹⁰ Moreover, we exclude self-employed workers, as they are usually not considered in calculating the gender pay gap. Altogether, we use a maximum of 143,014 observations. Out of these, 72,741 (50.9%) are male observations and 70,273 (49.1%) are female observations.

Our variable of interest is the (log) hourly wage for all respondents. As in module 2, ‘earnings’ and ‘income’ are used synonymously in this module, both referring to income earned on the labour market. However, EU-SILC provides only yearly gross earnings information including the following wage elements: wages and salaries paid in main and any secondary or casual job(s), supplementary payments (e.g. holiday payments, thirteenth-month payment), commissions and bonus payments. We follow the method of Berger and Schaffner (2015) to convert the yearly income information to hourly wages. Dividing the yearly income by 12 only provides a measure for the monthly wage if no job change or unemployment spells occur during the income reference period (the last calendar year or the last 12 months, depending on the country), which is the case for the majority of observations. Yearly income for persons with one unemployment spell and a single change in employment is obtained as a time-weighted average of income from the two jobs. Persons with more than two employment spells in the reference period cannot be considered in the study because the income information is based on main and second job. Finally, to derive hourly wages, we divide the monthly income by the sum of the usual working hours in the main and secondary job. We use the working hours from both jobs because the income variable refers to the salaries from the main and secondary job.

¹⁰ Student workers attending a higher ISCED-level (5–6) are not excluded as it is, like in the SES data, not possible to isolate student workers who do not engage at all in paid employment.

We use two different sets of explanatory variables. With the first set of variables we aim to replicate the SES-based analysis on EU-SILC data. We therefore use the same variables that are used in the gender pay gap calculation based on SES data. Namely, we use information on the age of the respondents (age 14–19, 20–29, 30–39, 40–49, 50–59, 60+) and the current employment status (full-time, high part-time, low part-time). The division into high and low part-time is performed based on information from the SES. We use the country average of the usual working hours of persons working more or less than 60% of a full-timer's normal hours in the local unit (in SES) as a threshold to divide self-defined part-time employment (in EU-SILC) into high part-time and low part-time. If the information for the specific country is missing in SES, we use the EU average in SES as a threshold for the respective country to specify part-time employment in EU-SILC. Furthermore, we include a dummy for the existence of a temporary contract. Education is classified according to the ISCED-scheme (ISCED 0–2, ISCED 3–4, ISCED 5–6). The occupation variable refers to the exercised occupation and is coded according to the international standard classification of occupations (ISCO). We use the information at the 2-digit-level (43 occupational groups). Industry-related information is provided by the 1-digit NACE classification. Two variables used in SES are not provided in EU-SILC; first, the dummy for public control on the enterprise and second, information on respondents' tenure.

The second set of variables explores the richness of variables in EU-SILC. With this dataset, we aim to explore the 'added value' of an EU-SILC based gender pay gap analysis compared to an SES-based one. While some variables are similar between the first and second set (employment status, ISCO, NACE), others are included in more detail. The age of the respondents is included as a metric variable. Education is now measured in six categories (ISCED 0, 1, 2, 3, 4, 5+6). Furthermore, the second set of variables includes additional variables, which are not provided in the SES data. We include a measure of overeducation, where overeducation is defined as closely as possible to the definition in module 1. The benchmark education in each occupational group is again the modal value of the education variable. However, two small discrepancies arise. First, the distinction of educational levels is less detailed in EU-SILC than in LFS (six educational groups instead of seven). Second, in EU-SILC, the variable identifying supervisory positions serves as an additional distinguishing feature capturing vertical segregation beyond the horizontal segregation that is measured with the ISCO classification.

The respondent's health status is measured in five pre-defined categories (very bad, bad, fair, good, very good) and information on the country of birth is used (current country of living, any European country except the current one or any other country). Furthermore, we use information on the respondent's marital status (married, not married). If information on the partner is available, we include the partner's income, employment status (employed, unemployed, disabled, inactive) and level of education (ISCED 0, 1, 2, 3, 4, 5+6). Firm size is defined by the number of employees and is measured on a scale of four categories (0–10, 11–19, 20–49, 50+). Finally, the existence of a supervisory position is included as a dummy.

Methodologically, all these variables are used as explanatory factors for decomposing the gender difference in log hourly mean wages for the total sample as well as for the gap at the country level. As a consequence, not only *decomposition results* will likely differ from those reported in module 2, but also the *size of the unadjusted pay gap* due to the outselection of observations with missing variable information. The econometric procedure is the same as in module 2: we first run wage regres-

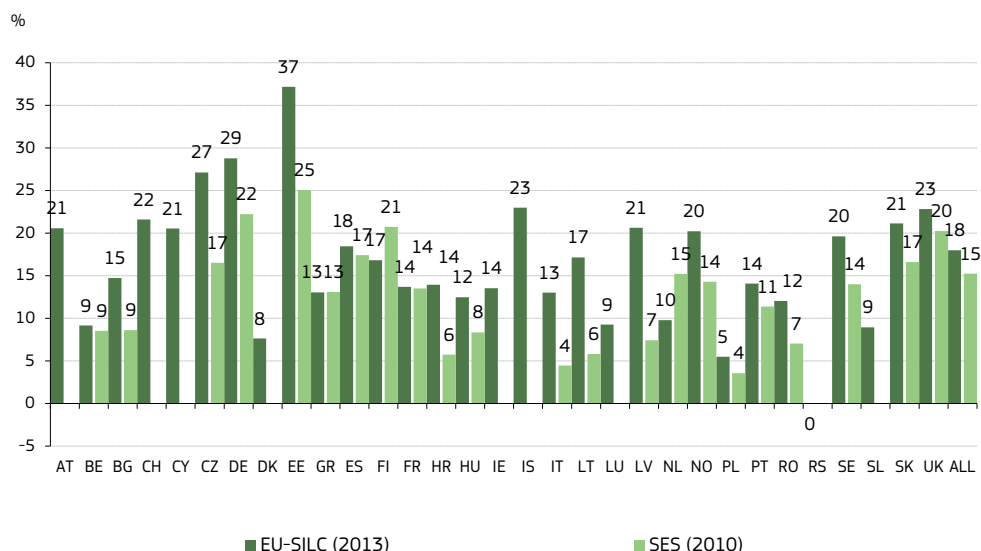
sions for men and women separately. Then, the coefficients obtained from the wage regressions are entered together with gender-specific averages of the explanatory factors the decomposition analysis. As a decomposition technique, we keep to the classical Oaxaca-Blinder decomposition to ensure comparability with module 2 (see the method section of module 2 as well as the *Methodological Guides* for a methodological description). In the following, the results of this approach are presented.

3. Results

3.1 Decomposition into explained and unexplained gap

Before analysing the decomposition results generated by our extended model, it is important to check for differences in the size of the unadjusted gender pay gaps between the EU-SILC and the SES databases. Figure 11 sheds light on these differences. In most countries, unadjusted pay gaps tend to be larger in the EU-SILC sample. This is likely due to the measurement discrepancies mentioned above, especially with regard to the wage definition. The size of the cross-country gap in EU-SILC is 18.4%, compared to 15.3% measured based on SES. Nevertheless, looking at the ranking of countries, a similar picture emerges. Most of those countries that exhibited high gaps under SES do the same under EU-SILC. Some exceptions to this might be worth noting. The most striking ones represent the Baltic countries, potentially because the comparatively small sample sizes reinforce the impact of selection issues. For Latvia and Lithuania, gaps around the EU average are measured by EU-SILC, while they were among the lowest in SES. Estonia is still the country with the largest overall pay gap in EU-SILC, but now with a considerably larger lead over the other countries. Moreover, Finland and the Netherlands stand out as the only countries with sizably smaller gaps under EU-SILC than under SES.

Figure 11: Size of the unadjusted gender pay gaps in EU-SILC and SES

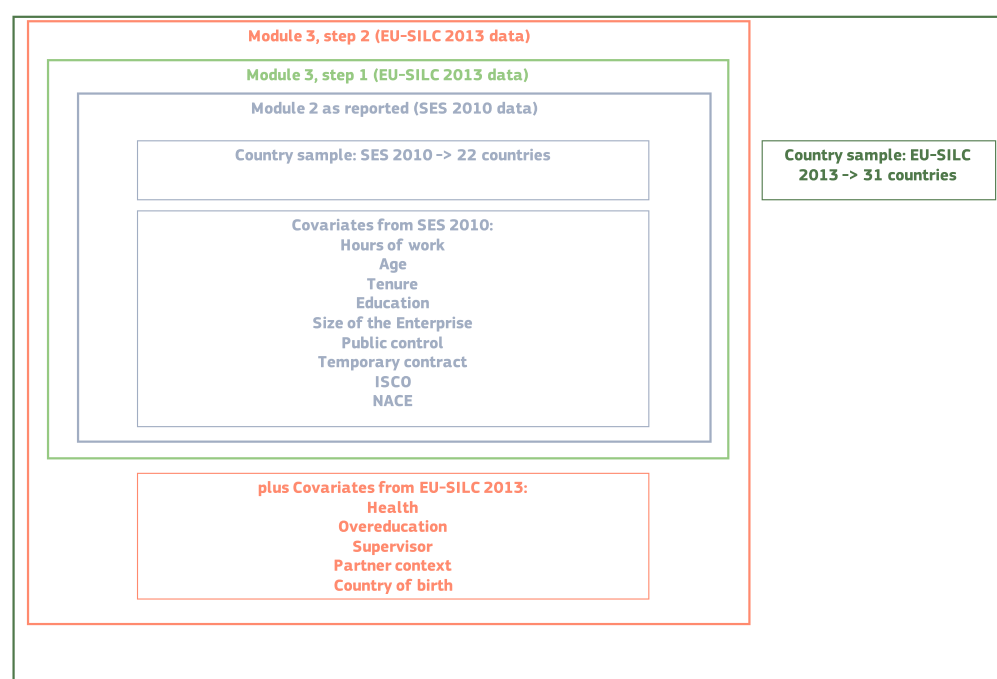


Sources: SES (2010), EU-SILC (2013), HWWI (2015).

The countries additionally incorporated in the EU-SILC based analysis tend to perform close to the cross-country average, with the exception of Denmark, Luxembourg, Serbia and Slovenia, which all exhibit gaps clearly below average. Serbia is a particularly interesting case, as it is the only country where the unadjusted gap is even (but only very slightly) negative. However, we can confirm our general statements on the geographical distribution of gender pay gaps formulated in module 2: unadjusted gaps tend to be smaller in the Eastern than in the Western part of Europe. At the same time, the correlation between pay gaps and female labour market participation at country level is reproduced in EU-SILC (see Figure A 1 in the Appendix). Therefore, in explaining the country differences, we still need to be aware of the fundamental measurement bias related to the employment decision (as outlined on page 79ff in module 2).

In the decomposition analysis, we applied different set-ups to reveal the impact of the changes made from modules 2 to 3. In more detail, we proceed in three steps (see Figure 12). We first tried to replicate the SES-country sample and model specification with EU-SILC data, as far as this was possible (i.e. excluding the variables job tenure and public control missing in EU-SILC). In a second step, we ran estimations based on an extended model including the additional variables available in EU-SILC, keeping the country sample constant. Finally, we applied the extended model to our enlarged dataset including 31 countries.

Figure 12: Estimation Procedure in Module 3



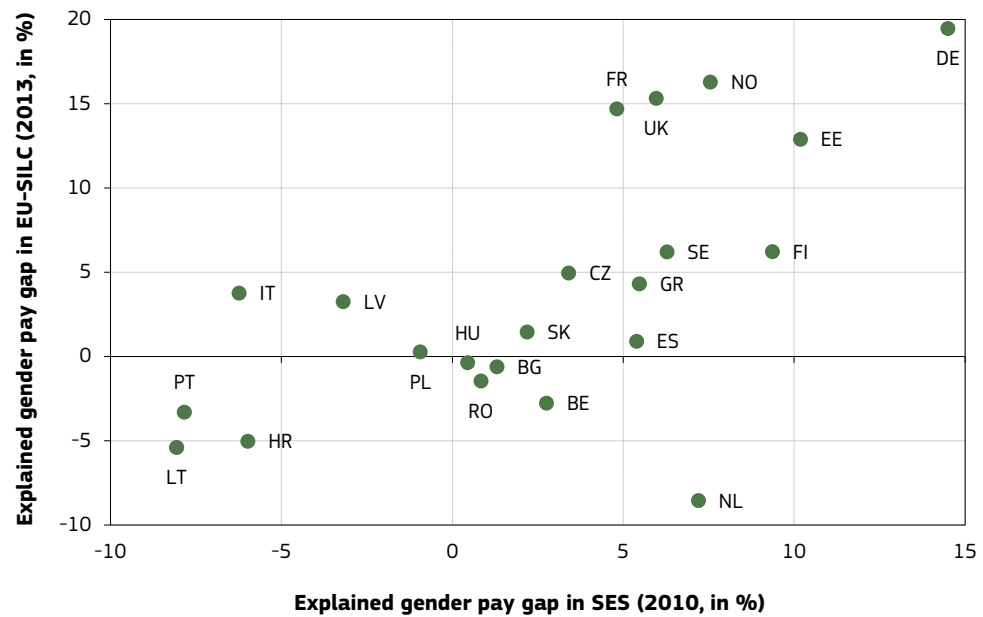
Source: HWWI (2015).

With the first estimation exercise, it appeared that replicating the SES results in EU-SILC did not yield a significant deviation from our previous decomposition results. The only decomposition variable that changed the sign of its characteristics effect is age (from positive to negative). However, the magnitude of the effect remains so

close to zero that this can hardly be considered a deviation. Results of the second step represent a mere subgroup of results obtained in the final step. Therefore, we restrict our detailed discussion to the most insightful estimation results, the decomposition of our extended model based on the enlarged country sample.

In analogy to module 2, we estimate and decompose the gap for an aggregate cross-country sample as well as for each country separately. The aggregate gap of 18.4% is divided into an explained part of 7.2% and an unexplained part of 11.1%. Hence, the share of explained wage variation has clearly increased compared to our previous results. Given the range of additional explanatory factors, this is the expected development. Nevertheless, the major part of gender wage differences is still left unexplained, stressing the role of unmeasured gender differences and evaluation effects.

Moreover, we are able to make country comparisons about the size of the explained gender pay gap. Figure 13 confronts the explained pay gaps from the last module with our new estimation results (omitting countries unavailable in module 2). It documents that the overall picture of country differences in the role of explanatory factors remains fairly stable. Again, we observe the by far largest explained gap for Germany. Norway, the UK, France and Estonia also exhibit explained gaps clearly above the cross-country average. In all of these countries except for Estonia, the unexplained gaps are at the same time considerably smaller than in cross-country average (see Table A 17 in the Appendix). This is also reminiscent of a similar pattern observed in SES. It shows that country differences in observed labour market characteristics are still insufficient to explain the ranking of unadjusted gender pay gaps among EU countries, even though we added several additional factors to the set of characteristics. The only country whose explained pay gap is dramatically affected by our change of model and dataset is the Netherlands. While their explained gap was measured to be positive and slightly above country average in SES (7.2%), it is now reduced to -8.6%, implying that the Netherlands exhibit the second smallest explained gap among all countries (after Slovenia). Apart from sample selection issues, this discrepancy is apparently closely tied to an interesting result for the hours of work variable (see discussion below). For Italy, an opposite, yet less dramatic development is observed. Its explained gap shifts from -6.7% to 3.2%, which comes very close to the EU-SILC average. The main deviation here concerns the impact of occupational selection.

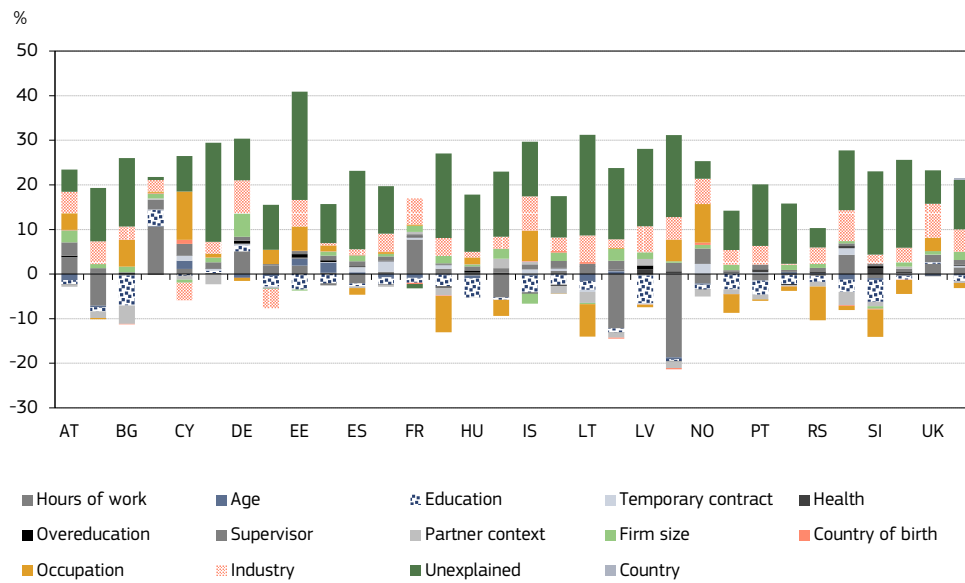
Figure 13: Explained gender pay gap according to SES and EU-SILC

Sources: EU-SILC (2013), SES (2010), HWWI (2015).

At the same time, the unexplained gender variation is smaller than 5% in the extended model for a range of countries (Austria, France, Norway, Serbia, Switzerland), something that was nowhere the case in module 2. In France, the unexplained gap is even negative (-1%), meaning that determinants other than observed endowment differences operate in favour of women. We will see that the constant term within the unexplained part bears major responsibility for this. In case of France it is slightly negative, indicating that French female workers outperformed their male counterparts in the endowment with statistically unobserved wage-determining characteristics and/or different remunerations to these unobservables.

3.2 Decomposition of the explained gender pay gap

Due to the extended set of explanatory factors, a further decomposition of explained and unexplained variation into the single characteristics and coefficients effects becomes even more informative. This holds first and foremost for the use of variables related to the partner context, which provide an insightful impression of the influence of household composition on gender wage differentials. In analogy to the representation chosen in module 2, Figure 14 depicts the contributions of variables (and variable groups) to the explained pay gaps at country level.

Figure 14: Decomposition of the gender pay gap (in %) in EU-SILC (2013)

Sources: EU-SILC (2013), HWWI (2015).

Comparison to SES

First of all, results for those factors of influence already included in module 2 are largely consistent with the previous estimations (see Figure 9). Gender differences in the sorting into *industries* (5.0%) and in *hours of work* (1.5%) again contribute most to the EU-wide gap. The result that women in Europe are under-represented in industries *ceteris paribus* (i.e. for given worker characteristics) high pay levels (and thus over-represented in industries with low pay levels) is confirmed. Likewise, the fact that female employees more often work part-time and receive a wage penalty per hour from this part-time work continues to hold for the cross-country sample. However, examining this second result, the degree of country heterogeneity is tremendously higher than in module 2. As mentioned above, the Netherlands stand out in particular with a contribution of –18.6% of gender differences in hours of work to the overall gap at country level. Further countries in which this contribution is considerably negative are Belgium, Ireland and Luxembourg. In all of these examples, the negative effect does not stem from an unexpected gender sorting, as also in these countries women are more likely to work part-time. Instead, it is due to the fact that part-time workers receive higher hourly wages than full-time workers according to the wage information stored in EU-SILC. Regarding the prevalence of part-time work, country differences in social attitudes towards part-time work, partly reflected in corresponding labour market legislation, are likely to be of great importance. For instance, according to Eurostat, the Netherlands in 2014 has been the country with the by far largest share of part-time workers in Europe¹¹. About every second Dutch worker worked part-time, as compared to about every fifth worker in EU average. This particularity is also manifested in Dutch law. According to Visser et al. (2004), part-time work is supported mainly by two kinds of laws: a prohibition of discrimination by working hours and the right of employees to adjust the working hours within

11 http://ec.europa.eu/eurostat/statistics-explained/index.php/Employment_statistics#Further_Eurostat_information

an existing work contract under certain conditions. The Dutch example shows that such legal support can effectively eliminate the wage-reducing effect of part-time work. Directly, it could have forced firms to pay part-time and full-time workers equally. Thereby, it could have raised the attractiveness of part-time work in general which might have entailed an increased productivity in part-time jobs enabling firms to pay even higher hourly wages in these jobs. These effects could have been responsible for the unusual observed effect of a wage premium for part-time work, even though the immediate legal effect by the anti-discrimination law is hard to disentangle from the technology effect postulated by Goldin (2014). The pay-attractive part-time work is particularly advantageous for female workers, as many women appreciate working reduced working hours for family reasons. Potentially, it is also attractive for men with high family-time preferences, which is increasingly observed in younger generations. Hence, in the Dutch example, the legislative strategy seems to have turned the gendered pattern of part-time work into a vehicle for mitigating the gender pay gap.

Moreover, in a cross-country comparison, the distribution of *temporary contracts* again widens the wage gap. Fixed-time contracts are more prevalent among female workers. Temporary employment is at the same time associated with lower hourly wages. However, the contribution to the wage differential merely amounts to 0.3 pp. In the large majority of countries, it is likewise positive and nowhere to be found in a significant negative range. In addition, the influence of gender differences in *schooling* reduces the gap by 1.3 pp, again consistent with previous results. It continues to hold that female workers possess higher average levels of education, which improves their position in the labour market. As in module 2, Germany represents an outlier in this; the schooling disadvantage of German female workers contributes 1.2 pp to the German pay gap. In Switzerland, which was introduced with module 3, the levels of schooling are even more unfavourably distributed from the perspective of women, accounting for 3.6 pp of the Swiss pay gap. For *occupational sorting*, the (at first sight) surprising negative contribution observed in module 2 is maintained. We can refer to our previous argumentation to qualify this result. Most importantly, we need to stress that our occupational classification still does not allow us to fully separate vertical from horizontal sorting and the fact that some low-skilled tasks are in part executed outside the formal labour market potentially biases our estimations. This is also reflected in the enormous degree of country heterogeneity, reaching from -8.2% in Croatia to 10.8% in Cyprus.

The only clear distinction between the datasets concerns the impact of *firm size*. Gender differences in the size of the firm were estimated to reduce the pay gap based on the SES sample, but they are estimated to increase it under our new set-up (by 1.8% in cross-country estimation). This also holds when we run our estimations in EU-SILC based on the country sample from SES. Its source is the changed distribution of men and women over firms with different size. In EU-SILC, women are measured to be over-represented in small firms instead of in large firms, implying that the wage bonus associated with large firms operate in favour of men. In addition to the different reference year (SES: 2010, EU-SILC: 2013), this discrepancy could stem from the different sample composition: in contrast to EU-SILC, in SES firms with fewer than 10 employees are only included for some countries. Hence, a strong presence of women in this subgroup of firms could explain the observed change in sign. At country level, the positive sign is remarkably stable, with only six countries exhibiting an opposite effect of firm size.

Additional characteristics: Personal context

In respect of the newly introduced variables, the outcome remains fairly intuitive. The consideration of gender differences in the state of *worker health* (measured by four categories) reveals a positive, but small influence (0.2%) on the gap. Male workers are on average healthier than their female counterparts. Better health, in turn, is associated with higher wages. In EU-SILC, the health variable refers to the self-assessment of respondents. We suggest that health masks some statistically unobserved individual traits that are associated with higher earnings, e.g. the internal locus of control (LOC). There is some evidence that self-rated health of persons who exhibit a high internal LOC is higher than that of persons with low LOC even when symptoms of physical illness are controlled for and that in gender comparison the effect of LOC matters more for men than for women (Pilisuk et al. 1993). That is, in our regression, the health effect on earnings (and subsequently, on the pay gap) might be due to a gender bias in self-reported health. Moreover, gender-specific responses to a worsening of the health condition could explain this pattern, as documented by a strong presence of men in physically demanding occupations where good health is a strict requirement. Exceptions at country level include Greece and Finland, however without causing a mitigation of the pay gap by a substantial degree.

As expected, gender differences in *supervisory positions* raise the gap, as men are over-represented in leading positions. The magnitude in cross-country analysis is quite considerable at 1.1%. It is also the only variable for which the sign of the characteristics effect is completely uniform across countries, indicating that supervisory power is a factor that contributes to the female earnings disadvantage everywhere in Europe. In Austria and Norway, its contribution is most striking, reaching levels higher than 3%. For the impact of *marital status*, gender differences widen the pay gap by 0.1%. For male workers, being married is associated with higher wages. At the same time, we observe more married men than women in our sample. We suggest that the lower labour market attachment of married women combined with the male breadwinner model is an important factor in shaping the marriage effect on the gender pay gap. In contrast, the impact of gender differences in the *country of birth* is practically zero.

Moreover, the introduction of *overeducation* as a determinant of the pay gap deserves attention in light of the previous results. In module 1, we found that at least among high-skilled workers, men are clearly more frequently overeducated than women. At the same time, empirical investigations have demonstrated that overeducation is associated with considerable wage penalties (Allen and Van der Velden, 2001; Bauer, 2002). However, we need to stress that comparability to the results of module 1 is limited. On the one hand, this is due to the use of different samples. The EU-SILC sample deviates from the LFS-samples both in considered countries and educational groups. On the other hand, measurement differences forced us to adjust our construction of the overeducation variable slightly. First, the distinction of educational levels is less detailed in EU-SILC than in LFS (six educational groups instead of seven). Second, in EU-SILC, the variable identifying supervisory positions serves as an additional distinguishing feature (besides the ISCO code) for defining occupational groups. It turns out that in EU-SILC, gender differences in overeducation status contribute to the gap. Overeducated workers receive lower average wages, a result that fits the existing literature well (see module 1 for more details). At the same time, the share of overeducated workers is higher among women than among men, which leads to the positive effect on the pay gap. The higher incidence for women than men contrasts with descriptive evidence from the sample of the

highly educated derived from EU-LFS in module 1. As mentioned, differences in the sample distribution in qualification and gender, and in the definition of overeducation, are likely to play a role in this. Nevertheless, the magnitude of the impact in the cross-country estimation remains very modest at 0.03%. At country level, the strongest effect is measured for Estonia (0.7%). At the same time, there are ten countries exhibiting negative estimates, including the Netherlands and the UK.

Additional characteristics: Partner context

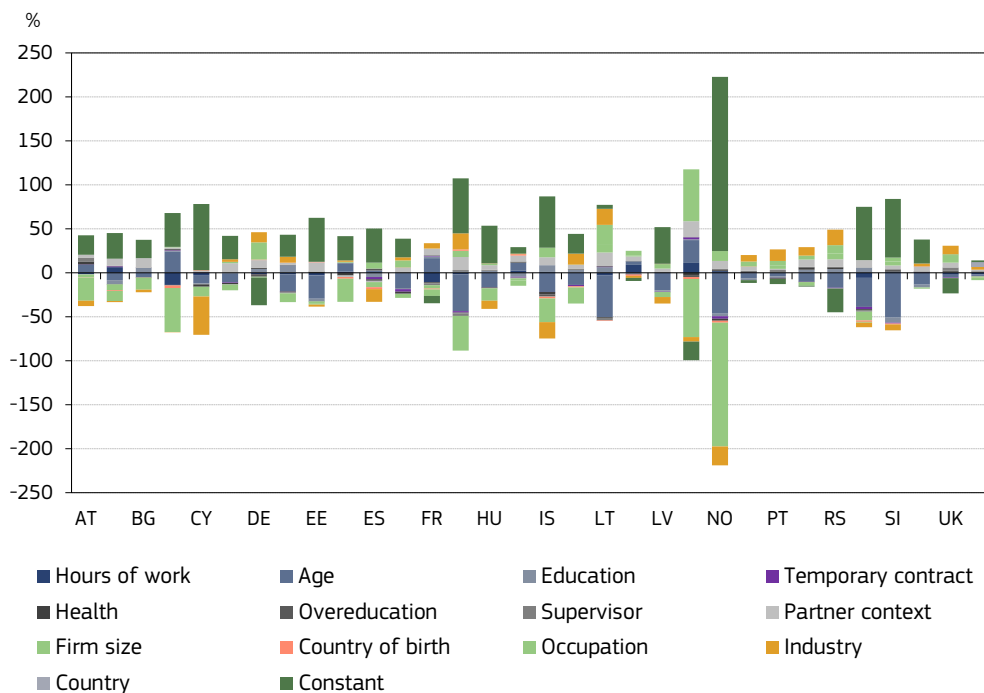
Furthermore, the exploitation of the partner information provided by EU-SILC yields interesting insights. It indicates an ambiguous role of partner matching. While the distribution of the educational levels of partners raises the gap, the distribution of partner income reduces it. Both outcomes could be attributed to economic similarity as a criterion for partner selection, a phenomenon well known in the literature as assortative mating: workers at the top end of the gender-specific distribution tend to match with workers at the top end of their gender-specific wage distribution and so forth (Mare, 1991; OECD, 2015). Hence, having a partner with relatively high income and high education signals a relatively high own labour income. Given that in absolute terms male workers earn higher average wages than female workers, female workers have on average the partner with the higher absolute income. This explains the negative contribution of the partner income to the explained part of the pay gap. Simultaneously, female workers possess the higher average education, in line with results for module 1. Hence, male workers have on average the partner with the higher educational level, implying that this effect operates in favour of men. In country comparison, these results prove remarkably stable. Exceptions where partner income has a positive effect on the pay gap are Hungary and Ireland.

Moreover, the distribution of inactive partners widens the gender gap by 0.5% in cross-country perspective. This is likely to result from a reverse influence: a high labour income reduces the partner's incentives to work. Hence, for men, having an inactive partner is a sign for a high own labour market reward. The fact that participation rates are lower among women than among men also reflects this partner income effect. In contrast, the distribution of unemployed partners mitigates the gap. Having an unemployed partner reduces the expected wage, potentially because it forces workers to accept bad jobs to keep the household afloat. This is also consistent with our results from module 1, where we found that living together with unemployed household members raises the risk of overeducation. As partner unemployment is more frequent among men's partners, the contribution to the gap is negative. However, the overall magnitude of the impact is considerably less impressive (-0.01%) than in case of inactive partners. Accordingly, comparing the single-country results, the positive sign for inactive partners proves to be clearly more stable than the negative sign for unemployed partners. Finally, for the distribution of disabled partners, an impact on the gender pay gap very close to zero is measured. This is not due to a potential irrelevance of this variable for the individual wage level. In fact, for both male and female workers, having a partner who is disabled is measured to be associated with higher wage levels. Instead, the reason is that the distribution of disabled partners shows no significant signs of gender asymmetry.

3.3 Decomposition of the unexplained part of the gender pay gap

As mentioned before, the unexplained gap consists of gender differences in the evaluation of the measured characteristics as well as of the influence of statistically unobserved variables captured by a constant term. In analysing results of the Oaxaca-Blinder-decomposition for the unexplained part, we witness considerable country heterogeneity in potentially discriminatory effects (see Figure 15). Nevertheless, some EU-wide patterns are detectable.

Figure 15: Decomposition of the unexplained gender pay gap (in %) in EU-SILC (2013)



Sources: EU-SILC (2013), HWWI (2015).

A difference in the evaluation of the partner context between male and female workers contributes to the pay gap in all countries except for Greece and Spain, with a net contribution of 1.6% in the aggregate sample. Having a partner with higher income entails a higher wage bonus for male than for female workers. In the context of assortative mating set out above, this reflects the notion of glass ceilings: the gender gap is larger at the top than at the bottom end of the wage distributions, since women still have trouble accessing leading positions (for evidence see e. g. Albrecht et al., 2003, 2004 and 2015). Hence, wage discrepancies between partners are likely to be larger for a high-income than for a low-income couple, implying that differences in partner income are statistically of smaller weight for the income of women. Furthermore and irrespective of women's position in the wage distribution, the finding that women's wages 'benefit' less from spouses' superior earnings than men points to the motivation by relative resources in combination with the traditional 'male breadwinner' model. In a bargaining perspective, men re-

act to a high earnings (bargaining) power of their female mates with catching up in own earnings whereas women are rather inclined to accept their additional earner position.¹² Moreover, the implications of having a partner who is not active on the labour market crucially differ by sex. While for men, inactive partners are associated with higher wages, the opposite holds for women. That is, additionally to the fact that men are more likely to be partnered to inactive persons, they also benefit more from the latter than women do. We argue that the motivation is quite similar in both cases, relying on the above cited male breadwinner model. While a woman's inactivity rather points to a prosperous male partner (i.e. having a male partner with high income reduces own incentives to work), inactivity of the male partner rather hints at bad income prospects for the household as a whole. The financial pressure that might force the female partner to take on any accessible job and/or again the influence of assortative mating (low-income couples) might represent the factors underlying this observation.

Moreover, the coefficients effect of industry is highly positive in the cross-country estimation (4.0%). In more detail, male workers in sectors other than the reference sector education received larger wage bonuses (compared to the reference sector) than female workers. That is, the gender pay inequity within sectors is more pronounced in the remaining sectors than in the benchmark sector education, pointing at higher costs of time flexibility in the remaining sectors.¹³ This reproduces our finding in module 2. Again, we can interpret this as a sign for a high degree of intra-sectoral sorting of male and female workers into different occupations and vertical positions; something we cannot fully control for with the industry and occupation variables available in EU-SILC. However, country heterogeneity concerning this impact is much more pronounced than in module 2. There exists a range of countries where the industry effect is measured to be highly negative, Norway with -21.6% representing an extreme case. Differences in the aggregation of sectors between the two datasets are likely to bear some responsibility for this, again limiting comparability.

Another feature whose wage impact differs significantly between men and women is age. As in case of module 2, the coefficients effect of age reduces the gender pay gap at the aggregate level (here by 2.1%). Examining the reasons for this, we need to be aware that we measure age now on a metric scale, not in terms of age groups as in module 2. Nevertheless, the coefficients of this metric age measure in the wage regressions hint at the expected explanations. They are positive for both male and female workers, but larger for females. Female workers thus receive a higher return to age than their male counterparts. This is consistent with our reasoning from module 2. We argued that this could first point at a sorting of women into less physically demanding occupations, where an increase in work experience with age is not overshadowed by a decline in physical productivity. Second, a stronger impact of age-dependent selection in case of female workers was discussed. With increasing age, women still have a stronger tendency to exit the labour market prematurely.

¹² Goldin (2014) shows that the partner income is particularly important in the combination with the existence of small children. Women with small children and high-income partners are more likely to quit the job since their own wage that is cut due to family compatible work arrangements seems not attractive enough to them to be worth remaining in employment.

¹³ According to the formula of the wage gap decomposition, the gender differences in the returns to a certain characteristic are weighted with the female endowment with the respective characteristic. Therefore, if in female-dominated sectors the sectoral wage bonus is higher for men than for women, a positive coefficients effect of industry will be measured.

If this primarily concerns less productive female workers, the remaining female employees are positively selected, explaining the statistical result of a shrinking gender pay gap within older age groups. Please note again (as discussed above in module 2) that the measured age effect is a *ceteris paribus* effect – conditional on all else being equal. That is, the single effect of age arising from our estimations measures the wage difference between women and men who exhibit the same qualification, experience, occupation and other wage-relevant characteristics. Therefore, this effect must not be confused with age-specific wage profiles for men and women which usually feature an increasing wage gap with increasing age. This is mainly due to the fact that in these cross-sectional illustrations, age simply carries (not displayed) other relevant factors like experience that notably differs between men and women.

Finally, we should draw attention to the role of the constant term within the unexplained pay gap. It further raises the cross-country gap by 2.1%. In general, it summarises the impact of all unmeasured wage determinants. Therefore, it is not surprising to see that its level has shrunk considerably in comparison to the previous module. Apparently, the additional personal and household variables introduced with EU-SILC did improve our understanding of the gender pay gap quite substantially. Of those factors still omitted, we should first mention actual work experience, for which we lack a sufficient proxy. Given women's backlog in this kind of experience, as caused by career breaks related to childbirth and parenting, a positive constant corresponds to basic intuition. Other unaddressed features could for example relate to gender differences in statistically unobserved abilities or negotiation skills. At the country level, again some exceptions are noticeable. We observe with Germany and the UK two large countries with negative constants of high magnitude. This demonstrates that any monocausal explanation not only finds its limits in the complexity of influencing factors, but also in the role of country particularities, e.g. in respect of the labour market and family institutions.

4. Conclusion

Reproducing the decomposition analysis of gender pay gaps in Europe based on EU-SILC data has proven to yield valuable additional insights. Making use of this dataset allowed us to add several potential determinants of gender wage differences in comparison to SES. At the same time, we were able to assess the stability of our results from module 2. With regard to the main impact factors, we found some confirmation for our previous argumentation. The sorting of male and female workers into industries with different pay levels is a major contributing factor to the gender pay gap in Europe. Women are over-represented in low-wage sectors such as health and social work activities, but under-represented in high-wage sectors like manufacturing. Gender differences in the amount of working hours, especially in the form that women work more often part-time, likewise account for the gender pay gap at the European level.

In a model comparison, the only clear deviation is marked by the impact of firm size, now predicting that the average female worker is disadvantaged compared to the average male worker by the fact that she works in a firm of different scale. However, this can be explained by the impact of data selection. As EU-SILC also contains firms with less than 10 employees, it accounts (unlike SES) for an over-representation of women in very small firms.

The characteristics effects of the personal context determinants introduced with EU-SILC are largely intuitive. Gender differences in worker health magnify the wage gap, just as the under-representation of women in supervisory positions does. Moreover, we found that the linkages between the pay gap and the phenomenon of overeducation are rather weak, even though we estimated a clearly wage-depressing effect for overeducation. This is because women in EU-SILC are only slightly more frequently overeducated than men. A worker's place of birth (domestic or abroad) is also of very low explanatory power for the pay gap.

Further information is obtained from the analysis of the partner context. Being married pays off more for male than for female workers and the latter are less often married than the former; both facts add to the pay gap. The fact that men and women have partners with different average income and educational attainment has important implications. While the distribution of the educational levels of partners raises the wage gap, the distribution of partner income reduces it. The phenomenon of assortative mating in connection with gender discrepancies in income and education is the likely key for understanding these results. They should therefore rather be interpreted as secondary expressions of the underlying gender differences in worker characteristics, which limits their usefulness for immediate policy conclusions. The partner's labour market status is also of relevance. The distribution of inactive partners amplifies gender wage differences, while the distribution of unemployed partners has a mitigating effect. This demonstrates the complexity of channels through which labour market policies can exert influence on the gender pay gap.

Finally, we need to stress again that we derived our conclusions merely from the subset of the population that was actually working. Persons outside the workfor-

ce can be expected to differ in major personal characteristics from those inside. Furthermore, as some of these characteristics are wage-relevant, their potential earnings will also differ. As the decision to go to work partly hinges upon expected payment, we can expect them to be lower than for actual workers on average, notwithstanding contrary effects for certain subgroups. Recent evidence confirms such a positive selection also for women (Jacobsen et al., 2015). This is a fact that leads us to assume that our analyses tend to underestimate the gender gap in potential hourly earnings. As female employment rates in most European countries are still lower than those of males, there exists a range of women that are discouraged from working by their low earnings prospects. These women disappeared from the statistics, making it hard to infer on the real extent of gender pay inequality. It is also likely to bias country comparisons, as revealed by the close correlation between national pay gaps and female employment rates. Therefore, an obvious task for future work will be to integrate inactive (but employable) persons into a more comprehensive empirical analysis of the roots of gender disparities in job opportunities. Such an analysis should ideally account for the interplay of wages and labour market activity as well as for the dynamics of career decisions over the lifecycle.

5. References – Modules 2 and 3

Aisenbrey, S. and Brückner, H. (2008): Occupational aspirations and the gender gap in wages. *European Sociological Review*, 24(5), 633–649.

Albrecht, J., Bjorklund, A. and Vroman, S. (2003): Is there a glass ceiling in Sweden? *Journal of Labor Economics*, 21(1), 145–177.

Albrecht, J., van Vuuren, A. and Vroman, S. (2004): Decomposing the gender wage gap in the Netherlands with sample selection adjustments. *IZA Discussion Paper Series No. 1400*.

Albrecht, J., Skogman Thoursie, P. and Vroman, S. (2015): Parental leave and the glass ceiling in Sweden (No. 2015: 4). *IFAU-Institute for Evaluation of Labour Market and Education Policy*.

Allen, S. (2001): Technology and the wage structure. *Journal of Labor Economics*, 19(2), 440–483.

Allen, J. and Van der Velden, R. (2001): Educational mismatches versus skill mismatches: Effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, 434–452.

Antonczyk, D., Fitzenberger, B. and Sommerfeld, K. (2010): Rising wage inequality, the decline of collective bargaining, and the gender wage gap. *Labour Economics*, 17(5), 835–847.

Arrow, K. (1973): The theory of discrimination. *Discrimination in labor markets*, 3(10), 3–33.

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2007): Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial & Labor Relations Review*, 60(2), 163–186.

Barnet-Verzat, C. and Wolff, F. C. (2008): Gender wage gap and the glass ceiling effect: a firm-level investigation. *International Journal of Manpower*, 29(6), 486–502.

Barón, J. D. and Cobb-Clark, D. A. (2010): Occupational segregation and the gender wage gap in private and public sector employment: A distributional analysis. *Economic Record*, 86(273), 227–246.

Bauer, T. K. (2002): Educational mismatch and wages: a panel analysis. *Economics of Education Review*, 21(3), 221–229.

Beblo, M., Behninger, D., Heinze, A. and Laisney, F. (2003): *Measuring Selectivity – Corrected Gender Wage Gaps in the EU*. ZEW Discussion Paper No. 3–74.

Becker, G. S. (1957): *The Economics of Discrimination*. Chicago (IL): University of Chicago Press.

Becker, G. S. (1965): A theory of the allocation of time. *The Economic Journal*, 75(299), 493–517.

Becker, G. S. (1981): *A Treatise on the Family*. Cambridge (MA): Harvard University Press.

Becker, G. S. (1985): Human capital, effort, and the sexual division of labour. *Journal of Labor Economics*, 3(1), 33–58.

Benenden Healthcare Society (2011): Study reveals different attitudes to sickness absence between men and women. Press Release from 25 May 2011. Download available under: <https://www.benenden.co.uk/newsroom/research/research-archive/study-reveals-different-attitudes-to-sickness-absence-between-men-and-women/>

Berger, M. and Schaffner, S. (2015): *A Note on How to Realise the Full Potential of the EU-SILC data*. ZEW Discussion Paper No. 15-005.

Bettio, F. (2002): The pros and cons of occupational gender segregation in Europe. *Canadian Public Policy/Analyse de Politiques*, 28 (supplement), 65–84.

Bettio, F. and Verashchagina, A. (2009): *Gender segregation in the labour market: root causes, implications and policy responses in Europe*. European Commission's Expert Group on Gender and Employment (EGGE), European Commission, Directorate-General for Employment, Social Affairs and Equal Opportunities.

Bittmann, M., England, P., Sayer, L., Folbre, N. and Matheson, G. (2003): When does gender trump money? Bargaining and time in household work. *American Journal of Sociology*, 109(1), 186–214.

Black, S. E. and Spitz-Oener, A. (2010): Explaining women's success: Technological change and the skill content of women's work. *The Review of Economics and Statistics*, 92(1), 187–194.

Blau, F. D. and Kahn, L. M. (1997): Swimming upstream: Trends in the gender wage differential in the 1980s. *Journal of Labor Economics*, 15(1), 1–42.

Blau, F. D. and Kahn L. M. (2003): Understanding international differences in the gender pay gap. *Journal of Labor Economics* 21(1), 106–144.

Blinder, A. S. (1973): Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436–455.

Boll, C., Hüning, H., Leppin, J. and Puckelwald, J. (2015): *Potential Effects of a Statutory Minimum Wage on the Gender Pay Gap – A Simulation-based Study for Germany*. Research Paper 163, HWWI: Hamburg.

Boll, C. and Leppin, J. S. (2015): Die geschlechtsspezifische lohnücke in Deutschland: Umfang, Ursachen und Interpretation. *Wirtschaftsdienst*, 95(4), 249–254.

Booth, A. L., Francesconi, M. and Frank, J. (2002): Temporary jobs: Stepping stones or dead ends? *The Economic Journal*, 112(480), 189–213.

Brines, J. (1994): Economic dependency, gender, and the division of labor at home. *American Journal of Sociology*, 100(3), 652–688.

Brown, J. N. (1989): Why do wages increase with tenure? On-the-job training and life cycle wage growth observed within firms. *The American Economic Review*, 97(5), 971–991.

Browning, M. and Chiappori, P. A. (1998): Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, 66(6), 1241–1278.

Brynin, M. and Francesconi, M. (2004): The material returns to partnership: the effects of educational matching on labour market outcomes and gender equality. *European Sociological Review*, 20(4), 363–377.

Busch, A. and Holst, E. (2009): *Glass Ceiling Effect and Earnings: The Gender Pay Gap in Managerial Positions in Germany*. Discussion paper No. 905: DIW German Institute for Economic Research, Berlin.

Daly, A., Meng, X., Kawaguchi, A. and Mumford, K. (2006): The gender wage gap in four countries. *Economic Record*, 82(257), 165–176.

Del Río, C., Gradín, C. and Cantó, O. (2011): The measurement of gender wage discrimination: The distributional approach revisited. *The Journal of Economic Inequality*, 9(1), 57–86.

DiNardo, J., Fortin, N. M. and Lemieux, T. (1996): Labour market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica*, 64(5), 1001–1044.

Eurostat-Data Explorer (2015): Part-time employment as percentage of the total employment, by sex and age (%) (lfsa_eppga), last update 31.08.2015. Extracted 23.09.2015 from:

<http://appsso.eurostat.ec.europa.eu/nui/show.do>.

European Commission, Directorate-General for Justice and Consumers (2015): Special Eurobarometer 428: Gender Equality – Report.

Firpo, S., Fortin, N. M. and Lemieux, T. (2007): *Unconditional Quantile Regressions*. Technical Working Paper 339, National Bureau of Economic Research.

Fuchs, V. R. (1989): Women's quest for economic equality. *The Journal of Economic Perspectives*, 3(1), 25–41.

Goldin, C. and Polachek, S. (1987): Residual differences by sex: Perspectives on the gender gap in earnings. *The American Economic Review*, 77(2), 143–151.

Goldin, C. (2014): A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4), 1091–1119.

Gornick, J. C. and Jacobs, J. A. (1998): Gender, the welfare state, and public employment: A comparative study of seven industrialized countries. *American Sociological Review*, 63(5), 688–710.

Gregory, R. G. and Borland, J. (1999): Recent developments in public sector labour markets. *Handbook of Labor Economics*, 3(C), 3573–3630.

Heckman, J. J. (1979): Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society*, 47(1), 153–161.

Henz, U. and Sundström, M. (2001): Partner choice and women's paid work in Sweden: The role of earnings. *European Sociological Review*, 295–316.

Hinz, T. and Gartner, H. (2005): Geschlechtsspezifische Lohnunterschiede in Branchen, Berufen und Betrieben/The gender wage gap within economic sectors, occupations, and firms. *Zeitschrift für Soziologie*, 34(1), 22–39.

Jacobsen, J., Khamis, M. and Yuksel, M. (2015): Convergences in men's and women's life patterns: Lifetime work, lifetime earnings, and human capital investment. In: S.W. Polachek, K. Tatsiramos and K. F. Zimmermann (eds.) *Gender Convergence in the Labour Market* (Vol. 41). Research in Labour Economics. Bingley, UK: Emerald Group Publishing.

Juhn, C., Murphy, K. M. and Pierce, B. (1993): Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3), 410–442.

Kalleberg, A. L. (2000): Nonstandard employment relations: Part-time, temporary and contract work. *Annual Review of Sociology*, 26, 341–365.

Kolberg, J. E. (1991): The welfare state as employer. London, NY: Sharpe.

Lewbel, A. (2007): Endogenous selection or treatment model estimation. *Journal of Econometrics*, 141(2), 777–806.

Light, A. and Ureta, M. (1995): Early-career work experience and gender wage differentials. *Journal of Labor Economics*, 13(1), 121–154.

Macpherson, D. A. and Hirsch, B. T. (1995): Wages and gender composition: Why do women's jobs pay less? *Journal of Labor Economics*, 13(3), 426–471.

Mandel, H. and Semyonov, M. (2005): Family policies, wage structures, and gender gaps: Sources of earnings inequality in 20 countries. *American Sociological Review*, 70(6), 949–967.

Manning, A. and Petrongolo, B. (2008): The part-time pay penalty for women in Britain. *The Economic Journal*, 118(526), F28–F51.

Manser, M. and Brown, M. (1980): Marriage and household decision-making: A bargaining analysis. *International Economic Review*, 21(1), 31–44.

Mare, R. D. (1991): Five decades of educational assortative mating. *American Sociological Review*, 56(1), 15–32.

Marini, M. M. (1989): Sex differences in earnings in the United States. *Annual Review of Sociology*, 15, 343–380.

Masters, S. H. (1969): An interindustry analysis of wages and plant size. *The Review of Economics and Statistics*, 51(3), 341–345.

McElroy, M. B. and Horney, M. J. (1981): Nash-bargained household decisions: Toward a generalisation of the theory of demand. *International Economic Review* 22(2), 333–349.

McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001): Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.

Meng, X. and Meurs, D. (2004): The gender earnings gap: Effects of institutions and firms—a comparative study of French and Australian private firms. *Oxford Economic Papers*, 56(2), 189–208.

Ministère du travail, de l'emploi, de la Formation Professionnelle et du dialogue social (2015): Ségrégation professionnelle et écarts de salaires femmes-hommes, DARES analyses, No. 082.

Munasinghe, L., Reif, T. and Henriques, A. (2008): Gender gap in wage returns to job tenure and experience. *Labour Economics*, 15(6), 1296–1316.

Oaxaca, R. (1973): Male–female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693–709.

OECD (2015): *In It Together: Why Less Inequality Benefits All*. Paris: OECD Publishing.

Oi, W.Y. and Idson, T. I. (1999): Firm size and wages. In: *Handbook of Labor Economics*, Vol. 3, Part B, Ch. 33, 2165–2214.

Petersen, T. and Morgan, L. A. (1995): Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *American Journal of Sociology*, 101(2), 329–365.

Pilisuk, M., Montgomery, M. B., Parks, S. H. and Acredolo, C. (1993): Locus of control, life stress, and social networks: Gender differences in the health status of the elderly. *Sex Roles* 28(3), 147–166.

Reskin, B. F. (2001): Labor markets as queues: A structural approach to changing occupational sex composition. *American Sociological Review*, 52(2), 195–210.

Reskin, B. F. and Roos, P. A. (2009): *Job queues, gender queues: Explaining women's inroads into male occupations*. Philadelphia, PA: Temple University Press.

Ruhm, C. J. (1998): The economic consequences of parental leave mandates: Lessons from Europe. *Quarterly Journal of Economics*, 113(1), 285–317.

Samuelson, P. A. (1956): Social indifference curves. *The Quarterly Journal of Economics*, 70(1), 1–22.

Shafer, E. F. (2011): Wives' relative wages, husbands' paid work hours, and wives' labour force exit. *Journal of Marriage and Family*, 73(1), 250–263.

Shannon, M. (1996): Minimum wages and the gender wage gap. *Applied Economics*, 28(12), 1567–1576.

Simón, H. (2012): The gender gap in earnings: An international comparison with European matched employer–employee data. *Applied Economics*, 44(15), 1985–1999.

Skirbekk, V. (2004): Age and individual productivity: A literature survey. *Vienna Yearbook of Population Research*, 2, 133–153.

Topel, R. (1991): Specific capital, mobility and wages: wages rise with job seniority. *The Journal of Political Economy* 99(1), 145–176.

Visser, J., Wilthagen, T., Beltzer, R. and van der Putte, E. (2004): Part-time employment in the Netherlands: from atypicality to a typicality. In: S. Sciarra, P. Davies and M. Freedland (eds.) *Employment Policy and the Regulation of Part-time Work in the European Union: A Comparative Analysis*. Cambridge, UK: Cambridge University Press.

Waldfogel, J. (1998): The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics*, 16(3), 505–545.

Wallerstein, M. (1999): Wage setting institutions and pay inequality in advanced industrial societies. *American Journal of Political Science*, 43(3), 649–680.

Weichselbaumer, D. and Winter-Ebmer, R. (2007): The effects of competition and equal treatment laws on gender wage differentials. *Economic Policy*, 22(50), 235–287.

Wood, R. G., Corcoran, M. E. and Courant, P. N. (1993): Pay differences among the highly paid: The male–female earnings gap in lawyers' salaries. *Journal of Labor Economics*, 11(3), 417–441.

Wright, E. O., Baxter, J. and Birkelund, G. E. (1995): The gender gap in workplace authority: A cross-national study. *American Sociological Review*, 60(3), 407–435.

SUGGESTIONS FOR ADDITIONAL STATISTICAL INDICATORS

In the course of our work, we faced several limitations associated with data constraints. Future research on gender issues in the context of labour markets could benefit from extending the accessed EU databases (LFS, SES, EU-SILC) in several dimensions. Based on our knowledge of existing research gaps, we would like to make the following propositions for relevant indicators to be included in one or more of the aforementioned databases:

- Actual work experience (in all three databases): It would be an important step forward to document the sum of years/months that a person has so far been in full-time or part-time employment (starting from labour market entry). This information is crucial for the investigation of wage penalties due to family-related breaks. With suitable data on employment biography, the unexplained part of the pay gap, amounting to 10.9% in the EU average according to our study, can be reduced in a notable manner, as research for Germany shows (Boll/Leppin 2015).

The indicator 'firm tenure' as years staying with the current employer (provided in the LFS data) is different from work experience, since it refers to employer mobility. Hence, 'tenure' information is suitable for calculating motherhood penalties on the firm level only, but not suitable for this aim at the country (or any other firm-overarching) level. Less satisfactory, but still a step forward would be 'potential work experience', calculated as current age minus years of education minus 6 (as a proxy for the entry age into education). The problem with this variable is that actual withdrawals from employment are not visible for the researcher. Hence, with potential experience, one of the most important criteria in which women differ from men (family breaks) and their wage effects cannot be correctly analysed (but only roughly estimated). We thus strongly prefer 'actual experience' over 'potential experience' and 'firm tenure'.

- Occupational position (in all three databases): More precise information on the actual job position of a person (both in the vertical and horizontal dimension) could give more clear-cut impressions of the extent of occupational segregation as well as of the existence of glass ceilings in career progression as two potential driving forces of gender inequality in the labour market.
- Number and age of children (in SES and EU-SILC): Information on family size and age structure of children is necessary for computing motherhood penalties. Moreover, it is also useful as an indicator of household productivity, thereby reflecting the reservation wage of women and hence their employment propensity.
- Job-related training and education (in SES and EU-SILC): Detailed information is required to investigate gender divides in participation and course duration as an important source of gender earnings divides in the life-course perspective. The information given in LFS is not suitable for two reasons; first, it refers to the 4 weeks preceding the interview only and second, the information cannot be linked to the income information and the household context information provided by SES and EU-SILC, respectively.

- Educational success (in all three databases): Additional information (apart from the level of the highest educational degree) on educational success would improve tracking gender differences in formal work skills and productivity signalling on the labour market. These could include the final grade received at graduation and/or the actual duration of university studies, and vocational training.

Appendix – Module 1

Table A 1: Descriptive statistics: high-skilled workers, EU-28

Variable	obs.	mean	std.	min.	max.
Overeducation	415197	0.3212	0.4669	0	1
Personal characteristics					
Female	416275	0.4985	0.5000	0	1
Age Group:					
- 15–24 years	416275	0.0325	0.1773	0	1
- 25–34 years	416275	0.2657	0.4417	0	1
- 35–44 years	416275	0.2983	0.4575	0	1
- 45–54 years	416275	0.2450	0.4301	0	1
- 55–64 years	416275	0.1400	0.3469	0	1
- 65–74 years	416275	0.0185	0.1348	0	1
Married	416275	0.5706	0.4950	0	1
Foreigner EU countries	416195	0.0366	0.1879	0	1
Foreigner non-EU countries	416195	0.0307	0.1725	0	1
Field of study:					
- Teaching, education	407486	0.0927	0.2900	0	1
- Humanities, languages, arts	407486	0.1003	0.3004	0	1
- Social sciences, business and law	407486	0.3112	0.4630	0	1
- Science, mathematics, computing	407486	0.0939	0.2917	0	1
- Engineering, manufacturing, construction	407486	0.1917	0.3937	0	1
- Agriculture, veterinary	407486	0.0215	0.1451	0	1
- Health, welfare	407486	0.1455	0.3526	0	1
- Services	407486	0.0431	0.2031	0	1
Household characteristics					
Number of unemployed adults in same household	337097	0.0563	0.2467	0	4
Number of inactive adults in same household	337097	0.2103	0.4907	0	6
Persons aged 75 or older in same household	416275	0.0006	0.0236	0	1
Number of children between 0 and 5 years in same household	337097	0.2579	0.5598	0	4
Number of children between 6 and 11 years in same household	337097	0.2316	0.5413	0	5
Number of children between 12 and 14 years in same household	337097	0.1053	0.3365	0	4
Job characteristics					
Firm size:					
- 0–10 persons	356010	0.2065	0.4048	0	1
- 11–19 persons	356010	0.0973	0.2964	0	1
- 20–49 persons	356010	0.1654	0.3715	0	1
- 50 and more persons	356010	0.5309	0.499	0	1

Industry:					
- Agriculture, forestry and fishing	414990	0.0118	0.1078	0	1
- Mining and quarrying	414990	0.0022	0.0469	0	1
- Manufacturing	414990	0.1050	0.3065	0	1
- Electricity, gas, steam and air conditioning supply	414990	0.0095	0.0968	0	1
- Water supply, sewerage, waste management and remediation activities	414990	0.0043	0.0654	0	1
- Construction	414990	0.0372	0.1892	0	1
- Wholesale and retail trade; repair of motor vehicles and motorcycles	414990	0.0885	0.2840	0	1
- Transportation and storage	414990	0.0255	0.1578	0	1
- Accommodation and food service activities	414990	0.0196	0.1386	0	1
- Information and communication	414990	0.0547	0.2274	0	1
- Financial and insurance activities	414990	0.0468	0.2112	0	1
- Real estate activities	414990	0.0093	0.0958	0	1
- Professional, scientific and technical activities	414990	0.1093	0.3120	0	1
- Administrative and support service activities	414990	0.0278	0.1643	0	1
- Public administration and defence; compulsory social security	414990	0.0899	0.2860	0	1
- Education	414990	0.1556	0.3625	0	1
- Human health and social work activities	414990	0.1540	0.3610	0	1
- Arts, entertainment and recreation	414990	0.0206	0.1421	0	1
- Other service activities	414990	0.0225	0.1483	0	1
- Activities of households as employers; undifferentiated goods and service-producing activities of households for own use	414990	0.0041	0.0642	0	1
- Activities of extraterritorial organisations and bodies	414990	0.0019	0.0432	0	1
Marginal employment	406910	0.0303	0.1713	0	1
Temporary contract	351643	0.0971	0.2961	0	1
Usual working hours (in 10h)	406910	3.8347	1.0890	0.1	8
Usual working hours squared (in 10h)	406910	15.891	8.4739	0.01	64
Tenure (in 10 y)	412751	1.0440	0.9634	0	6
Tenure squared (in 10 y)	412751	2.0182	3.2534	0	32
Participation in lifelong learning	415687	0.1612	0.3677	0	1
Second job	416237	0.0484	0.2146	0	1
Degree of urbanisation	416275	1.6792	0.7756	1	3

Sources: EU-LFS (2013), HWWI (2015).

Table A 2: Descriptive statistics: medium-skilled workers, EU-28

Variable	obs.	mean	std.	min.	max.
Overeducation	691106	0.1893	0.3918	0	1
Personal characteristics					
Female	693017	0.4468	0.4972	0	1
Age Group:					
- 15–24 years	693017	0.0705	0.2561	0	1
- 25–34 years	693017	0.2157	0.4113	0	1
- 35–44 years	693017	0.2683	0.4431	0	1
- 45–54 years	693017	0.2836	0.4507	0	1
- 55–64 years	693017	0.1468	0.3539	0	1
- 65–74 years	693017	0.0151	0.1220	0	1
Married	693017	0.5576	0.4967	0	1
Foreigner EU countries	692888	0.0301	0.1707	0	1
Foreigner non-EU countries	692888	0.0250	0.1562	0	1
Field of study:					
- Teaching, education	666137	0.0218	0.146	0	1
- Humanities, languages, arts	666137	0.0321	0.1761	0	1
- Social sciences, business and law	666137	0.2651	0.4414	0	1
- Science, mathematics, computing	666137	0.0275	0.1636	0	1
- Engineering, manufacturing, construction	666137	0.3910	0.488	0	1
- Agriculture, veterinary	666137	0.0397	0.1953	0	1
- Health, welfare	666137	0.0999	0.2999	0	1
- Services	666137	0.1229	0.3283	0	1
Household characteristics					
Number of unemployed adults in same household	583596	0.0666	0.2708	0	5
Number of inactive adults in same household	583596	0.2541	0.5371	0	7
Persons aged 75 or older in same household	693017	0.0006	0.0246	0	1
Number of children between 0 and 5 years in same household	583596	0.1999	0.4919	0	5
Number of children between 6 and 11 years in same household	583596	0.2145	0.5086	0	5
Number of children between 12 and 14 years in same household	583596	0.1151	0.3448	0	4
Job characteristics					
Firm size:					
- 0–10 persons	595038	0.2713	0.4446	0	1
- 11–19 persons	595038	0.1193	0.3241	0	1
- 20–49 persons	595038	0.1647	0.3709	0	1
- 50 and more persons	595038	0.4448	0.4969	0	1
Industry:					
- Agriculture, forestry and fishing	691513	0.0345	0.1825	0	1
- Mining and quarrying	691513	0.0034	0.0585	0	1
- Manufacturing	691513	0.1802	0.3844	0	1
- Electricity, gas, steam and air conditioning supply	691513	0.0096	0.0974	0	1

- Water supply, sewerage, waste management and remediation activities	691513	0.0078	0.0878	0	1
- Construction	691513	0.0834	0.2764	0	1
- Wholesale and retail trade; repair of motor vehicles and motorcycles	691513	0.1541	0.361	0	1
- Transportation and storage	691513	0.0579	0.2335	0	1
- Accommodation and food service activities	691513	0.0420	0.2006	0	1
- Information and communication	691513	0.0227	0.1488	0	1
- Financial and insurance activities	691513	0.0305	0.172	0	1
- Real estate activities	691513	0.0093	0.0958	0	1
- Professional, scientific and technical activities	691513	0.0400	0.1959	0	1
- Administrative and support service activities	691513	0.0395	0.1948	0	1
- Public administration and defence; compulsory social security	691513	0.0714	0.2575	0	1
- Education	691513	0.0429	0.2025	0	1
- Human health and social work activities	691513	0.1199	0.3249	0	1
- Arts, entertainment and recreation	691513	0.0137	0.1163	0	1
- Other service activities	691513	0.0288	0.1674	0	1
- Activities of households as employers; undifferentiated goods- and service-producing activities of households for own use	691513	0.0078	0.0879	0	1
- Activities of extraterritorial organisations and bodies	691513	0.0007	0.0258	0	1
Marginal employment	675874	0.0395	0.1949	0	1
Temporary contract	589890	0.0862	0.2806	0	1
Usual working hours (in 10h)	675874	3.7331	1.0988	0.1	8
Usual working hours squared (in 10h)	675874	15.1433	8.3185	0.01	64
Tenure (in 10y)	682450	1.0949	1.005	0	6
Tenure squared (in 10 y)	682450	2.2088	3.4956	0	32
Participation in lifelong learning	692262	0.0956	0.294	0	1
Second job	692971	0.0371	0.1891	0	1
Degree of urbanisation	693017	1.9497	0.8032	1	3

Sources: EU-LFS (2013), HWWI (2015).

Table A 3: Estimation results high-skilled workers: Austria, Belgium, Bulgaria, Cyprus, and Czech Republic

	Probability of overeducation: marginal effects				
	High-skilled workers aged 15–74				
	AT	BE	BG	CY	CZ
Personal characteristics					
Female	0.055	-0.026	-0.098*	0.117***	0.152**
Age groups (reference: 25–34 years):					
- 15–24 years	0.269***	0.004	0.154	0.496***	0.415***
- 35–44 years	-0.017	0.020*	0.035	-0.035	-0.000
- 45–54 years	0.028	0.017	0.077	-0.066**	-0.099**
- 55–64 years	0.068**	0.057***	0.007	-0.131***	-0.036
- 65–74 years	0.053	0.106	0.154	-0.146**	-0.083
Married	0.020	-0.008	-0.002	0.022	-0.008
Foreigner: EU countries	0.071***	0.061***		0.178***	0.031
Foreigner: non-EU countries	0.127***	0.139***		0.193***	0.060
Field of study (reference: social sciences, business and law):					
- Teaching, education	-0.143***	0.011	0.170**	0.115**	-0.027
- Humanities, languages, arts	-0.017	0.092***	0.202**	0.106**	0.184*
- Science, mathematics, computing	-0.014	0.072***	0.006	-0.017	0.043
- Engineering, manufacturing, construction	0.210***	0.064***	0.031	0.076***	0.137***
- Agriculture, veterinary	0.170***	0.191***	0.043	0.129	0.088
- Health, welfare	-0.247***	0.011	-0.011	-0.012	0.001
- Services	0.129***	0.078***	0.234***	0.343***	0.324***
Household characteristics					
Number of unemployed adults ¹	0.067	0.028**	0.049	0.055***	0.058
Number of inactive persons ¹	0.007	0.019***	0.008	0.027**	-0.027
Persons aged 75 or older ¹	-0.079				
Number of children between 0 and 5 years ¹	-0.010	0.005	-0.011	-0.016	-0.012
Number of children between 6 and 11 years ¹	-0.020	0.006	-0.006	-0.033	-0.014
Number of children between 12 and 14 years ¹	-0.050**	0.021**	-0.049	0.056**	0.022
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	-0.068***	-0.013	-0.076***	-0.086***	-0.039
- 20–49 persons	-0.095***	-0.026***	-0.096***	-0.102***	-0.088**
- 50 and more persons	-0.121***	-0.036***	-0.065**	-0.079***	-0.047
Marginal employment ²	0.106**	-0.014	-0.139**	-0.222***	-0.259***
Temporary contract	0.009	0.036***	-0.076	-0.027	-0.011
Usual working hours	0.018	-0.036***	-0.443***	-0.120***	-0.010
Usual working hours squared	-0.007*	0.004**	0.067***	0.009**	-0.010
Tenure	0.009	-0.005	-0.061*	-0.119***	-0.055
Tenure squared	-0.005	-0.005	0.020*	0.023***	0.009

Participation in LLL	-0.056***	-0.003	-0.010	-0.021	-0.042
Second job	-0.042**	0.008	-0.082	0.172***	-0.054
Degree of urbanisation	-0.001	-0.010*	0.039**	-0.005	-0.005
Interaction terms:					
Sex and married	-0.012	0.016	0.025	0.008	0.001
Sex and urbanisation	-0.014	-0.004	0.017	0.006	0.009
Sex and elder household members					
Sex and children:					
- number of children (0–5 years)	0.016	-0.021**	-0.033	0.023	-0.018
- number of children (6–11 years)	0.000	-0.011	0.034	0.088***	0.046
- number of children (12–14 years)	-0.028	-0.026*	0.139**	0.004	-0.034
Sex and age groups (references: 25–34 years):					
- 15–24 years	-0.067	0.008	0.070	-0.228***	-0.252***
- 35–44 years	-0.029	0.005	-0.039	-0.052	-0.090
- 45–54 years	-0.011	0.019	-0.100**	0.040	-0.021
- 55–64 years	0.057	-0.005	-0.049	0.125*	0.015
- 65–74 years	0.325***	-0.009	0.084	-0.204***	0.170
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	-0.060	0.011	-0.074	-0.162***	-0.157***
- Humanities, languages, arts	-0.062	-0.022	-0.082*	-0.141***	-0.190***
- Science, mathematics, computing	0.231***	0.027	0.118	-0.072*	-0.125*
- Engineering, manufacturing, construction	-0.179***	0.001	0.086	-0.190***	-0.179***
- Agriculture, veterinary	0.011	0.017	0.014	0.043	-0.057
- Health, welfare	0.124***	0.004	-0.036	-0.118***	-0.091
- Services	0.108	-0.012	-0.028	-0.086**	-0.155*
Observations	12,520	11,818	2,360	5,854	2,610

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 4: Estimation results high-skilled workers: Germany, Denmark, Estonia, Spain, and Finland

	Probability of overeducation: marginal effects				
	High-skilled workers aged 15–74				
	<i>DE</i>	<i>DK</i>	<i>EE</i>	<i>ES</i>	<i>FI</i>
Personal characteristics					
Female	-0.006	0.078***	-0.043	-0.070*	-0.024
Age groups (reference: 25–34 years):					
- 15–24 years	0.129***	0.261***	-0.076***	0.069	0.128
- 35–44 years	0.073***	0.053***	-0.001	0.014	0.056**
- 45–54 years	0.079***	0.047**	0.024	0.045	0.105***
- 55–64 years	0.091***	0.040*	0.038	0.008	0.104***
- 65–74 years	0.153***	0.013	0.089	0.087	0.008
Married	-0.006	0.006	-0.001	0.001	-0.014
Foreigner: EU countries	0.101***	0.097***		0.097**	0.043
Foreigner: non-EU countries	0.149***	0.222***	0.136***	0.198***	0.258**
Field of study (reference: social sciences, business and law):					
- Teaching, education	-0.230***	-0.060***	-0.195***	-0.029	0.029
- Humanities, languages, arts	-0.100***	0.061**	0.021	0.052	0.035
- Science, mathematics, computing	-0.088***	0.047*	0.122	0.024	0.077*
- Engineering, manufacturing, construction	-0.184***	0.125***	0.050	0.140***	-0.053***
- Agriculture, veterinary	0.072***	0.108**	-0.012	0.111**	0.053
- Health, welfare	-0.050***	-0.008	-0.071	0.000	-0.022
- Services	0.066***	0.532***	0.011	0.075	0.159**
Household characteristics					
Number of unemployed adults ¹	0.052***		0.018	0.039***	
Number of inactive persons ¹	-0.008		-0.032**	0.028***	
Persons aged 75 or older ¹	0.436***				
Number of children between 0 and 5 years ¹	-0.012		-0.010	0.007	
Number of children between 6 and 11 years ¹	-0.014*		-0.042*	-0.014	
Number of children between 12 and 14 years ¹	-0.002		0.021	-0.006	
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	0.008	-0.107***	0.031	-0.010	-0.011
- 20–49 persons	-0.054***	-0.100***	0.019	-0.055***	-0.042***
- 50 and more persons	-0.061***	-0.129***	0.029*	-0.019	-0.045***
Marginal employment ²	0.091***	-0.059	-0.047	-0.057	-0.039
Temporary contract	0.045***	0.095***	0.048	0.076***	0.117***
Usual working hours	-0.036***	-0.196***	-0.065**	-0.070**	-0.049
Usual working hours squared	-0.000	0.020***	0.010***	0.005	0.006
Tenure	0.033***	-0.019	-0.047***	0.020	0.023

Tenure squared	-0.013***	-0.004	0.006	-0.007	-0.013**
Participation in LLL	-0.051***	-0.034***	-0.049***	-0.034***	-0.013
Second job	-0.006	-0.009	0.035	-0.013	0.011
Degree of urbanisation	-0.002	0.011*	-0.011	0.022**	0.005
Interaction terms:					
Sex and married	0.011	-0.005	-0.016	-0.003	0.011
Sex and urbanisation	0.053***	-0.010	-0.011	-0.007	0.011
Sex and elder household members	-0.259				
Sex and children:					
- number of children (0–5 years)	-0.039***		-0.017	-0.041*	
- number of children (6–11 years)	-0.022**		0.033	0.027	
- number of children (12–14 years)	0.000		-0.014	0.031	
Sex and age groups (references:					
25–34 years):					
- 15–24 years	0.117***	-0.126***	0.180	0.011	-0.037
- 35–44 years	-0.016	0.016	0.003	0.021	-0.023
- 45–54 years	0.021	0.086***	-0.007	-0.025	-0.044*
- 55–64 years	0.013	0.116***	-0.011	0.015	-0.037
- 65–74 years	-0.109**	0.182**	-0.017	-0.077	0.106
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	0.070***	-0.112***	0.964***	0.096*	-0.061*
- Humanities, languages, arts	0.014	0.007	0.025	0.034	-0.028
- Science, mathematics, computing	-0.033	-0.050	-0.017	0.098*	0.023
- Engineering, manufacturing, construction	0.013	0.037	0.092*	-0.018	0.077*
- Agriculture, veterinary	-0.071**	-0.040	0.367***	0.151	0.124
- Health, welfare	0.024	-0.159***	0.069	0.061	-0.052
- Services	0.083***	-0.176***	-0.016	0.041	-0.010
Observations	50,363	16,069	3,253	10,115	3,898

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 5: Estimation results high-skilled workers: France, Greece, Croatia, Hungary, and Ireland

	Probability of overeducation: marginal effects				
	High-skilled workers aged 15–74				
	FR	GR	HR	HU	IE
Personal characteristics					
Female	0.052	0.003	0.103	0.041	-0.009
Age groups (reference: 25–34 years):					
- 15–24 years	0.284***	0.035	0.052	0.100**	0.004
- 35–44 years	0.038	0.020**	0.060	0.084***	0.006
- 45–54 years	0.008	0.013	0.085	0.105***	0.003
- 55–64 years	0.049	0.009	0.173**	0.092***	0.013
- 65–74 years	0.261	0.063	0.253*	0.143**	-0.009
Married	-0.031	-0.001	-0.078	-0.040**	-0.004
Foreigner: EU countries	0.098	0.131***	0.037	-0.075	0.061***
Foreigner: non-EU countries	0.148**	0.243***			0.028***
Field of study (reference: social sciences, business and law):					
- Teaching, education	-0.222***	-0.038***	-0.073	0.039	-0.015*
- Humanities, languages, arts	-0.005	0.003	-0.029	-0.063*	0.005
- Science, mathematics, computing	0.055	0.030**	0.166**	0.065**	0.059***
- Engineering, manufacturing, construction	0.049*	0.063***	0.140***	-0.001	0.053***
- Agriculture, veterinary	0.163**	0.134***	0.313***	0.064**	0.148***
- Health, welfare	-0.310***	0.123***	-0.123	0.101**	0.022**
- Services	0.095*	-0.007	0.379***	0.316***	0.004
Household characteristics					
Number of unemployed adults ¹	0.027	0.012***	0.044	0.015	0.015***
Number of inactive persons ¹	0.028*	0.010***	0.028*	0.012*	0.005*
Persons aged 75 or older ¹					
Number of children between 0 and 5 years ¹	-0.024	-0.005	0.080	-0.015	0.001
Number of children between 6 and 11 years ¹	-0.026	-0.005	0.067	-0.024*	-0.001
Number of children between 12 and 14 years ¹	0.019	-0.004	-0.087*	-0.025	-0.007
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	-0.045**	-0.003	0.087*	-0.038***	-0.008*
- 20–49 persons	-0.049**	-0.013**	-0.076**	-0.079***	-0.005
- 50 and more persons	-0.042**	0.032***	-0.068**	-0.062***	0.008**
Marginal employment ²	-0.165***	-0.021**	-0.109	-0.182***	0.011
Temporary contract	0.033	0.014*	0.188***	0.088***	0.011*
Usual working hours	-0.050	0.016	0.140	-0.167***	0.002
Usual working hours squared	-0.003	-0.002	-0.018	0.018***	0.000
Tenure	0.033	-0.018**	0.013	-0.045***	-0.012**
Tenure squared	-0.021***	0.004	0.004	0.007	0.001
Participation in LLL	-0.033**	-0.029***	-0.139	-0.043**	-0.013**

Second job	-0.020	0.031	0.198**	-0.019	0.045***
Degree of urbanisation	0.021*	0.010***	0.045*	0.014*	0.008***
Interaction terms:					
Sex and married	0.011	-0.022**	0.130*	-0.005	0.006
Sex and urbanisation	0.005	-0.029***	0.016	0.005	-0.013***
Sex and elder household members					
Sex and children:					
- number of children (0–5 years)	-0.002	0.010	-0.141**	-0.014	-0.007*
- number of children (6–11 years)	0.029	0.016**	-0.072	0.018	-0.005
- number of children (12–14 years)	-0.017	-0.001	0.084	0.042*	0.006
Sex and age groups (references: 25–34 years):					
- 15–24 years	-0.111**	-0.009	-0.044	-0.008	-0.018*
- 35–44 years	-0.055	0.045**	-0.073	-0.063***	0.007
- 45–54 years	-0.052	0.039**	-0.090	-0.115***	-0.003
- 55–64 years	-0.034	0.141***	-0.246***	-0.095***	-0.017**
- 65–74 years	-0.072	0.113	0.286	-0.103**	0.060
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	0.055	0.063**	-0.178***	-0.068***	0.003
- Humanities, languages, arts	-0.018	-0.015	-0.132	0.191**	-0.000
- Science, mathematics, computing	-0.031	-0.016	-0.208***	0.039	0.010
- Engineering, manufacturing, construction	-0.127***	-0.012	-0.213***	-0.011	0.001
- Agriculture, veterinary	-0.060	-0.032***	-0.261***	0.048	-0.030***
- Health, welfare	0.161*	-0.031***	-0.164**	0.048	-0.011
- Services	0.095	0.060*	-0.097	-0.094***	-0.017**
Observations	7,871	13,517	2,302	14,675	23,947

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 6: Estimation results high-skilled workers: Italy, Lithuania, Luxembourg, Latvia, and the Netherlands

	Probability of overeducation: marginal effects				
	High-skilled workers aged 15–74				
	<i>IT</i>	<i>LT</i>	<i>LU</i>	<i>LV</i>	<i>NL</i>
Personal characteristics					
Female	0.110***	0.027	-0.022	-0.096**	-0.013
Age groups (reference: 25–34 years):					
- 15–24 years	-0.025	-0.022	-0.097	0.038	0.020
- 35–44 years	-0.033*	0.030	0.073	0.061*	-0.043*
- 45–54 years	-0.071***	0.036	0.039	0.069*	-0.019
- 55–64 years	-0.154***	0.053*	-0.083	0.112**	-0.028
- 65–74 years	-0.238***	0.178***	0.308	0.220**	-0.069
Married	-0.001	-0.047**	0.036	-0.096***	-0.039**
Foreigner: EU countries	0.148***	-0.089*	-0.059**		-0.059**
Foreigner: non-EU countries	0.431***	0.418***	-0.008	0.106***	0.011
Field of study (reference: social sciences, business and law):					
- Teaching, education	-0.016	0.001	-0.147**	-0.011	-0.008
- Humanities, languages, arts	0.061**	0.069	0.076	0.065	0.031
- Science, mathematics, computing	0.049**	0.021	0.050	0.007	0.062**
- Engineering, manufacturing, construction	-0.117***	0.068***	0.075	0.036	0.072***
- Agriculture, veterinary	0.071	0.013	0.252	0.127*	0.150**
- Health, welfare	-0.283***	0.122*	0.065	-0.046	0.048
- Services	-0.044	0.126**	0.492***	0.091*	0.146***
Household characteristics					
Number of unemployed adults ¹	-0.003	0.056***	0.121	-0.038	0.060**
Number of inactive persons ¹	0.020***	0.023***	0.007	0.029***	0.038**
Persons aged 75 or older ¹	-0.022	0.146			
Number of children between 0 and 5 years ¹	-0.014	-0.003	-0.014	0.059***	-0.013
Number of children between 6 and 11 years ¹	-0.017	0.021	-0.051*	-0.033	-0.018
Number of children between 12 and 14 years ¹	-0.043*	0.007	-0.043	-0.039	-0.011
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	-0.034**	-0.002	-0.071**	-0.015	-0.011
- 20–49 persons	-0.095***	0.032*	-0.051	-0.034**	-0.043**
- 50 and more persons	-0.073***	0.034***	-0.070*	-0.032*	-0.044**
Marginal employment ²	0.152***	-0.018	-0.034		-0.016
Temporary contract	-0.006	0.161***	0.036	0.076	0.042**
Usual working hours	0.271***	-0.067	-0.061	-0.002	-0.110***
Usual working hours squared	-0.037***	0.014	0.007	0.004	0.001
Tenure	-0.033*	-0.110***	-0.101**	-0.047**	0.026
Tenure squared	-0.001	0.022***	0.028*	0.002	-0.011**

Participation in LLL	-0.062***	-0.047***	-0.050*	-0.056***	0.002
Second job	-0.005	-0.035***	-0.072	0.015	-0.008
Degree of urbanisation	-0.003	0.038***	-0.062***	-0.020*	0.011
Interaction terms:					
Sex and married	-0.023	0.033	-0.054	0.073**	0.002
Sex and urbanisation	-0.006	-0.030***	0.016	0.011	0.002
Sex and elder household members	0.405				
Sex and children:					
- number of children (0–5 years)	0.037**	-0.016	0.042	-0.072***	-0.035*
- number of children (6–11 years)	0.035*	0.014	0.068*	0.059**	-0.003
- number of children (12–14 years)	0.054*	0.030	0.068	0.038	-0.066***
Sex and age groups (references: 25–34 years):					
- 15–24 years	0.070	0.040	0.253	0.047	0.090
- 35–44 years	-0.009	0.002	0.001	-0.045	0.067*
- 45–54 years	0.005	-0.018	0.058	-0.071***	0.029
- 55–64 years	0.046	-0.032	0.181	-0.048*	0.017
- 65–74 years	-0.144	-0.018		-0.026	0.081
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	-0.172***	-0.055*	0.257	-0.025	-0.034
- Humanities, languages, arts	-0.078***	-0.066***	-0.011	0.016	-0.013
- Science, mathematics, computing	-0.123***	-0.057**	-0.110**	0.134	0.018
- Engineering, manufacturing, construction	-0.020	-0.000	-0.112*	0.043	0.094**
- Agriculture, veterinary	-0.045	0.086		0.146	-0.082
- Health, welfare	-0.104***	-0.101***	-0.113**	-0.052	-0.027
- Services	-0.093*	-0.038	0.010	-0.067***	-0.017
Observations	25,516	8,695	2,539	3,610	11,292

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 7: Estimation results high-skilled workers: Portugal, Romania, Sweden, Slovakia, and the United Kingdom

	Probability of overeducation: marginal effects				
	High-skilled workers aged 15–74				
	<i>PT</i>	<i>RO</i>	<i>SE</i>	<i>SK</i>	<i>UK</i>
Personal characteristics					
Female	0.030	-0.045**	0.033***	0.018	0.007
Age groups (reference: 25–34 years):					
- 15–24 years	0.019	0.093**	0.091***	0.382***	0.017
- 35–44 years	0.008	-0.053***	0.024***	-0.001	-0.029
- 45–54 years	0.045	-0.074***	0.063***	-0.023	-0.044
- 55–64 years	-0.002	-0.077***	0.075***	-0.011	0.005
- 65–74 years	0.191*	-0.009	0.016	-0.061	0.085
Married	-0.086***	0.017	0.008	-0.015	-0.089***
Foreigner: EU countries	0.122	0.850***	0.103***	-0.066	0.156**
Foreigner: non-EU countries	0.327***	-0.112***	0.190***	0.266**	0.067
Field of study (reference: social sciences, business and law):					
- Teaching, education	-0.004	-0.086**	-0.043***	0.022	0.011
- Humanities, languages, arts	0.059	-0.057**	0.102***	0.263***	0.055*
- Science, mathematics, computing	0.127***	-0.080***	0.220***	0.012	-0.039
- Engineering, manufacturing, construction	-0.019	-0.067***	0.027***	0.063*	0.035
- Agriculture, veterinary	-0.080	-0.033	0.077***	0.186***	0.022
- Health, welfare	-0.174***	-0.063***	0.026**	-0.000	-0.147***
- Services	0.104**	0.071***	-0.024**	0.115**	0.219***
Household characteristics					
Number of unemployed adults ¹	-0.027	0.074***		0.052**	0.069**
Number of inactive persons ¹	-0.002	0.017***		0.001	-0.001
Persons aged 75 or older ¹	-0.084				
Number of children between 0 and 5 years ¹	-0.076***	0.012		0.018	-0.015
Number of children between 6 and 11 years ¹	0.024	0.023		0.015	0.005
Number of children between 12 and 14 years ¹	0.008	-0.009		-0.019	-0.005
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	-0.044*	-0.048***	-0.048***	-0.003	0.020
- 20–49 persons	-0.031	-0.080***	-0.079***	-0.171***	-0.021
- 50 and more persons	0.022	-0.035***	-0.070***	-0.116***	-0.057***
Marginal employment ²	-0.093		-0.063***	-0.003	-0.171***
Temporary contract	0.087***	0.042	0.113***	0.195***	-0.025
Usual working hours	-0.114***	0.106	-0.059***	0.001	-0.211***
Usual working hours squared	0.014***	-0.008	0.001	-0.008	0.016***
Tenure	-0.063**	-0.117***	0.034***	-0.087***	-0.066***
Tenure squared	0.014**	0.031***	-0.015***	0.028***	0.009
Participation in LLL	-0.019	0.059*	-0.033***	-0.032	-0.051***

Second job	-0.012	0.133***	0.010	-0.000	0.082**
Degree of urbanisation	-0.003	0.042***	-0.004	-0.009	0.013
Interaction terms:	0.071**	0.021	-0.031***	0.007	0.088***
Sex and married	-0.004	-0.022**	0.003	0.057***	0.000
Sex and urbanisation	0.149				
Sex and elder household members					
Sex and children:					
- number of children (0–5 years)	0.083***	-0.049***		0.013	-0.065**
- number of children (6–11 years)	-0.043	0.001		0.103***	-0.005
- number of children (12–14 years)	-0.038	0.057**		0.059	0.081**
Sex and age groups (references: 25–34 years):					
- 15–24 years	0.069	0.026	-0.007	-0.311***	0.092
- 35–44 years	-0.072**	0.004	-0.009	0.007	-0.021
- 45–54 years	-0.111***	0.026	-0.010	-0.014	0.035
- 55–64 years	-0.143***	-0.000	0.004	-0.120**	-0.045
- 65–74 years	-0.218***	-0.001	-0.014	-0.129	-0.095
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	-0.110***	-0.014	-0.076***	-0.225***	-0.088
- Humanities, languages, arts	-0.055	0.102**	-0.053***	-0.268***	-0.076*
- Science, mathematics, computing	-0.073	0.125***	-0.060***	-0.048	-0.082**
- Engineering, manufacturing, construction	0.013	0.085***	-0.019*	0.003	-0.093*
- Agriculture, veterinary	-0.007	0.039	-0.000	-0.166**	-0.049
- Health, welfare	0.056	-0.014	-0.094***	-0.024	-0.041
- Services	0.072	0.093**	0.203***	0.337***	-0.101
Observations	8,881	12,548	46,628	6,214	6,987

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 8: Estimation results medium-skilled workers: Austria, Belgium, Bulgaria, Cyprus, and Czech Republic

	Probability of overeducation: marginal effects				
	Medium-skilled workers aged 15–74				
	AT	BE	BG	CY	CZ
Personal characteristics					
Female	0.056***	-0.009	0.005	0.134***	-0.009
Age groups (reference: 25–34 years):					
- 15–24 years	-0.027**	0.010	-0.015	0.025	-0.007*
- 35–44 years	-0.005	-0.008	0.015	-0.062***	0.000
- 45–54 years	-0.014	-0.026***	0.026	-0.018	-0.006
- 55–64 years	0.001	-0.019	0.031	-0.035	-0.007*
- 65–74 years	-0.028	0.004	0.155	-0.033	0.011
Married	0.013*	-0.010	0.003	-0.003	-0.000
Foreigner: EU countries	0.067***	0.018		0.070***	0.006
Foreigner: non-EU countries	0.059***	0.058*		0.207***	0.007
Field of study (reference: social sciences, business and law):					
- Teaching, education	0.047	0.017	0.493*	0.846***	0.010
- Humanities, languages, arts	0.001	0.042*	-0.017	0.082**	0.012
- Science, mathematics, computing	0.137***	0.003	-0.018	-0.022	0.010
- Engineering, manufacturing, construction	0.035***	-0.002	-0.076**	0.032	-0.036***
- Agriculture, veterinary	0.074***	0.053**	-0.028***	-0.069	-0.015***
- Health, welfare	-0.016	-0.024	-0.002	0.232**	0.107**
- Services	-0.059***	-0.002	-0.013	0.408***	-0.009***
Household characteristics					
Number of unemployed adults ¹	-0.002	0.014	0.002	-0.016	0.007**
Number of inactive persons ¹	-0.006	-0.001	-0.003	-0.013	-0.004
Persons aged 75 or older ¹	-0.037				
Number of children between 0 and 5 years ¹	0.010	0.006	0.005	-0.022	0.006*
Number of children between 6 and 11 years ¹	-0.013**	0.001	0.009	-0.018	-0.002
Number of children between 12 and 14 years ¹	0.006	0.005	0.015	-0.028	0.004
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	0.003	-0.007	-0.004	-0.015	0.000
- 20–49 persons	0.016**	-0.018***	-0.005	-0.026*	-0.003
- 50 and more persons	0.038***	-0.022***	-0.003	0.003	-0.002
Marginal employment ²	0.012	-0.001	-0.025***	-0.140***	0.031
Temporary contract	0.003	0.018	0.024	-0.015	0.005
Usual working hours	-0.052***	-0.025**	-0.075***	-0.061**	0.010
Usual working hours squared	0.006***	0.002	0.009***	0.004	-0.001
Tenure	-0.018**	0.012	0.006	0.020	-0.001
Tenure squared	0.002	-0.007***	-0.001	-0.010	0.000
Participation in LLL	0.018**	0.010		0.035	0.007*

Second job	0.024**	-0.005	0.331*	0.007	-0.000
Degree of urbanisation	-0.031***	0.011**	-0.007	-0.014	-0.004**
Interaction terms:					
Sex and married	0.004	0.007	-0.011	0.008	-0.002
Sex and urbanisation	0.006	-0.018**	0.006	0.029**	0.002
Sex and elder household members	-0.005				
Sex and children:					
- number of children (0–5 years)	-0.023**	-0.014	-0.043*	0.001	-0.008*
- number of children (6–11 years)	0.009	-0.005	-0.008	0.029	-0.001
- number of children (12–14 years)	-0.014	-0.008	0.004	0.024	-0.011*
Sex and age groups (references: 25–34 years):					
- 15–24 years	0.025	0.040*	0.006	0.019	0.008
- 35–44 years	-0.018	-0.008	-0.011	-0.003	-0.000
- 45–54 years	-0.010	-0.002	-0.024***	-0.086***	-0.004
- 55–64 years	-0.034**	-0.001	-0.026***	-0.067**	0.001
- 65–74 years	-0.033	0.126	-0.026***	0.115	-0.009
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	0.016	0.021	-0.024**	-0.179***	0.028
- Humanities, languages, arts	-0.019	0.012	0.028	-0.089***	-0.002
- Science, mathematics, computing	0.099	-0.028	0.127	-0.012	0.014
- Engineering, manufacturing, construction	0.053***	0.031	0.023	-0.050*	0.008
- Agriculture, veterinary	0.020	0.068	0.223	0.591***	0.065
- Health, welfare	-0.013	0.076**	0.255	-0.073	-0.004
- Services	0.141***	0.040	0.023	-0.070***	0.014
Observations	47,952	14,119	3,618	6,176	10,277

1: in same household

2: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 9: Estimation results medium-skilled workers: Germany, Denmark, Estonia, Spain, and Finland

	Probability of overeducation: marginal effects				
	Medium-skilled workers aged 15–74				
	DE	DK	EE	ES	FI
Personal characteristics					
Female	-0.007	0.000	-0.112**	0.005	-0.033
Age groups (reference: 25–34 years):					
- 15–24 years	-0.047***	0.008	-0.010	0.052	-0.044***
- 35–44 years	-0.008	0.016	-0.079***	-0.013	0.051***
- 45–54 years	-0.018***	0.045***	0.040	0.017	0.047***
- 55–64 years	-0.017***	0.042***	0.095***	-0.100	0.044**
- 65–74 years	0.073***	0.058*	-0.003	-0.371***	-0.026***
Married	-0.000	-0.011	0.071***	0.002	0.012**
Foreigner: EU countries	0.073***	0.021	-0.011	0.187***	-0.004
Foreigner: non-EU countries	0.009	0.156***	0.042***	0.084	0.049
Field of study (reference: social sciences, business and law):					
- Teaching, education	0.101***	0.017	-0.133***	-0.043	-0.026***
- Humanities, languages, arts	-0.064***	-0.027	-0.025	-0.040	-0.009
- Science, mathematics, computing	-0.055***	-0.038***	-0.044	-0.158***	-0.021***
- Engineering, manufacturing, construction	-0.159***	-0.008	-0.060	0.206***	-0.035***
- Agriculture, veterinary	-0.074***	-0.004	-0.052	0.302***	-0.022***
- Health, welfare	0.120***	-0.088***	-0.151***	-0.045	-0.026***
- Services	-0.081***	0.277***	-0.034	0.281***	-0.016**
Household characteristics					
Number of unemployed adults ¹	-0.030***		0.018	0.053***	
Number of inactive persons ¹	-0.015***		0.006	0.032**	
Persons aged 75 or older ¹	0.000				
Number of children between 0 and 5 years ¹	0.003		0.017	0.017	
Number of children between 6 and 11 years ¹	-0.000		-0.020	-0.006	
Number of children between 12 and 14 years ¹	0.001		-0.023	-0.049	
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	0.017***	-0.004	0.003	0.035	0.003
- 20–49 persons	0.012***	0.006	0.001	-0.035	-0.004
- 50 and more persons	0.043***	0.028***	0.004	-0.002	-0.004
Marginal employment ²	0.014*	0.035	-0.104***	-0.110	-0.019***
Temporary contract	-0.001	0.020	0.025	0.050*	-0.008
Usual working hours	0.028***	-0.036**	-0.090**	-0.039	-0.005
Usual working hours squared	-0.002**	0.004**	0.009*	0.002	-0.000
Tenure	0.005	-0.037***	-0.030	-0.062	0.006
Tenure squared	-0.006***	0.004*	0.000	-0.000	-0.004**
Participation in LLL	0.055***	0.000	-0.046**	-0.059*	-0.001

Second job	0.014**	0.001	0.047	0.028	-0.003
Degree of urbanisation	-0.024***	-0.008**	-0.043***	0.039**	0.004
Interaction terms:					
Sex and married	-0.001	-0.008	-0.094***	0.013	-0.012**
Sex and urbanisation	-0.004	-0.002	0.025**	-0.041*	0.002
Sex and elder household members	-0.036				
Sex and children:					
- number of children (0–5 years)	0.017***		-0.011	-0.053	
- number of children (6–11 years)	0.005		0.032	0.007	
- number of children (12–14 years)	0.002		0.004	0.131**	
Sex and age groups (references: 25–34 years):					
- 15–24 years	0.033***	-0.044**	-0.007	-0.106	0.839***
- 35–44 years	-0.029***	-0.015	0.177***	-0.026	0.051
- 45–54 years	-0.030***	-0.037***	0.100*	-0.056	0.033
- 55–64 years	-0.063***	0.001	0.077	0.030	0.034
- 65–74 years	-0.115***	-0.056*	0.231**	0.475***	0.974***
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	0.124***	0.065	0.909***	0.178	0.975***
- Humanities, languages, arts	0.122***	0.235***	-0.088***	0.196*	-0.020***
- Science, mathematics, computing	0.088***	-0.023	0.253	0.187**	0.081
- Engineering, manufacturing, construction	0.084***	0.325***	0.147**	0.032	0.022
- Agriculture, veterinary	0.132***	0.051	0.032	0.073	0.044
- Health, welfare	0.025**	0.091	0.766***	0.132	-0.007
- Services	0.038***	-0.087***	0.041	0.150*	-0.008
Observations	121,103	17,598	4,358	5,565	4,891

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 10: Estimation results medium-skilled workers: France, Greece, Croatia, Hungary, and Ireland

	Probability of overeducation: marginal effects				
	Medium-skilled workers aged 15–74				
	FR	GR	HR	HU	IE
Personal characteristics					
Female	0.032	-0.072**	0.017	-0.042***	-0.038***
Age groups (reference: 25–34 years):					
- 15–24 years	0.011	0.038	-0.064***	-0.014*	-0.002
- 35–44 years	-0.017	0.015	-0.005	-0.013**	0.012**
- 45–54 years	-0.049***	-0.028	0.007	-0.022***	-0.003
- 55–64 years	-0.045***	-0.008	0.052**	-0.027***	0.002
- 65–74 years	-0.083***	-0.138***	0.109	-0.028	-0.040***
Married	0.002	0.023	-0.005	-0.010*	-0.018***
Foreigner: EU countries	0.023	0.294***		-0.057***	0.071***
Foreigner: non-EU countries	0.016	0.198***			-0.005
Field of study (reference: social sciences, business and law):					
- Teaching, education	-0.082***	-0.064*	0.020	0.135	0.044
- Humanities, languages, arts	-0.048***	0.007	0.003	0.017	0.068***
- Science, mathematics, computing	-0.003	0.106**	0.095*	-0.020	0.038***
- Engineering, manufacturing, construction	-0.033***	0.323***	-0.092***	-0.012	0.010
- Agriculture, veterinary	-0.012	0.260***	-0.029**	-0.038***	0.161***
- Health, welfare	-0.086***	0.138***	-0.063***	-0.047***	-0.032***
- Services	0.036*	0.157***	-0.047***	-0.034***	-0.002
Household characteristics					
Number of unemployed adults ¹	-0.004	0.040***	-0.011	0.002	0.013**
Number of inactive persons ¹	-0.001	0.001	-0.002	-0.002	0.008**
Persons aged 75 or older ¹			-0.059***		0.123
Number of children between 0 and 5 years ¹	-0.008	0.002	-0.009	0.003	0.010***
Number of children between 6 and 11 years ¹	-0.012*	-0.010	-0.001	-0.002	-0.000
Number of children between 12 and 14 years ¹	0.004	0.007	-0.050***	-0.013*	-0.001
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	0.016	-0.027**	0.039***	0.013**	0.000
- 20–49 persons	-0.000	-0.030**	0.038***	0.019***	0.014**
- 50 and more persons	0.014*	0.008	0.019**	0.016***	0.016***
Marginal employment ²	-0.027	-0.044	0.057	-0.003	0.011
Temporary contract	-0.006	0.083***	0.047***	0.072***	0.021**
Usual working hours	0.010	0.012	0.003	0.006	-0.016*
Usual working hours squared	-0.002	-0.002	-0.003	-0.004*	0.000
Tenure	0.010	0.023	0.008	-0.034***	0.002
Tenure squared	-0.009***	-0.016**	-0.001	0.003*	-0.004*
Participation in LLL	0.034***	-0.104***	-0.040**	0.029*	-0.018

Second job	0.001	-0.029	0.028	-0.024**	0.075***
Degree of urbanisation	-0.002	0.032***	-0.010*	-0.002	0.009***
Interaction terms:					
Sex and married	-0.020*	-0.031	0.001	-0.014**	0.011
Sex and urbanisation	-0.010	-0.026**	-0.003	0.016***	-0.008*
Sex and elder household members			0.934***		
Sex and children:					
- number of children (0–5 years)	0.015	-0.005	0.006	-0.018**	-0.031***
- number of children (6–11 years)	0.021**	0.016	-0.016	0.004	-0.014**
- number of children (12–14 years)	-0.001	-0.018	0.030	0.002	0.009
Sex and age groups (references: 25–34 years):					
- 15–24 years	-0.015	0.074	0.074	-0.026***	-0.019
- 35–44 years	0.014	0.024	-0.033**	0.027**	-0.016*
- 45–54 years	-0.007	0.131***	-0.063***	0.023*	-0.023***
- 55–64 years	0.030	0.150**	-0.053***	0.058***	-0.008
- 65–74 years	0.913***			0.049	0.004
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	0.901***	0.135	-0.003	-0.053**	-0.009
- Humanities, languages, arts	-0.036*	0.060	0.071	0.022	-0.021*
- Science, mathematics, computing	-0.019	-0.011	-0.056***	0.000	0.008
- Engineering, manufacturing, construction	-0.020	-0.101***	0.062*	0.094***	0.109***
- Agriculture, veterinary	0.036	-0.039	-0.006	0.114***	-0.029**
- Health, welfare	0.051	-0.070***	0.148	0.085**	0.032
- Services	-0.014	0.234***	-0.016	0.058***	-0.012
Observations	13,916	9,571	6,199	39,090	14,347

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 11: Estimation results medium-skilled workers: Italy, Lithuania, Luxembourg, Latvia and the Netherlands

	Probability of overeducation: marginal effects				
	Medium-skilled workers aged 15–74				
	<i>IT</i>	<i>LT</i>	<i>LU</i>	<i>LV</i>	<i>NL</i>
Personal characteristics					
Female	-0.145***	-0.039	0.099	-0.017	0.036*
Age groups (reference: 25–34 years):					
- 15–24 years	0.033**	0.000	-0.038	-0.024	-0.025*
- 35–44 years	0.002	0.031	0.003	0.008	0.001
- 45–54 years	-0.011	0.061**	0.011	0.004	0.028*
- 55–64 years	-0.015	0.066**	-0.027	0.044	0.025
- 65–74 years	-0.113**	0.012	0.445**	-0.008	-0.023
Married	-0.018*	0.034*	0.036	-0.008	-0.001
Foreigner: EU countries	0.476***	0.112	0.018	0.178	0.017
Foreigner: non-EU countries	0.471***	0.088	0.181*	0.030	0.011
Field of study (reference: social sciences, business and law):					
- Teaching, education	0.223*	0.165	-0.048	0.439	0.069
- Humanities, languages, arts	0.219***	-0.096***	0.120	-0.021	-0.009
- Science, mathematics, computing	0.149***	-0.046	-0.006	-0.042	0.005
- Engineering, manufacturing, construction	0.124***	-0.095***	0.030	-0.122**	-0.026***
- Agriculture, veterinary	0.157***	-0.043	-0.037	-0.045	-0.044***
- Health, welfare	0.139***	-0.011	-0.072**	-0.022	0.008
- Services	0.201***	-0.121***	0.154**	-0.082**	-0.014
Household characteristics					
Number of unemployed adults ¹	0.051***	0.002	-0.082**	0.008	-0.018
Number of inactive persons ¹	0.026***	-0.001	-0.000	0.010	0.006
Persons aged 75 or older ¹	-0.051				
Number of children between 0 and 5 years ¹	-0.004	-0.022	0.009	-0.054**	-0.003
Number of children between 6 and 11 years ¹	0.013*	-0.000	-0.014	-0.068***	0.007
Number of children between 12 and 14 years ¹	0.002	0.007	-0.013	0.047	-0.012
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	-0.042***	0.038	-0.004	-0.009	0.017
- 20–49 persons	-0.074***	0.067***	0.018	-0.036*	0.008
- 50 and more persons	-0.073***	0.084***	0.004	-0.034*	0.016**
Marginal employment ²	0.154***	0.150	0.018	-0.096**	0.025
Temporary contract	0.073***	0.043	-0.059**	-0.014	-0.015**
Usual working hours	0.037**	-0.043	0.098***	0.030	-0.037***
Usual working hours squared	-0.006**	0.012	-0.011**	-0.003	0.005**
Tenure	-0.003	-0.022	-0.088***	0.012	-0.017**
Tenure squared	-0.022***	-0.003	0.027***	-0.020**	0.003
Participation in LLL	-0.060***	-0.042*	-0.022	-0.041	-0.008

Second job	-0.092***	-0.035**	0.021	0.069	0.022*
Degree of urbanisation	0.038***	-0.010	-0.012	-0.012	-0.001
Interaction terms:					
Sex and married	-0.012	-0.021	-0.055*	0.011	-0.011
Sex and urbanisation	0.003	0.002	-0.045*	-0.007	-0.010
Sex and elder household members					
Sex and children:					
- number of children (0–5 years)	0.004	-0.003	0.052*	0.049	0.004
- number of children (6–11 years)	-0.014	0.027	0.018	0.006	-0.012
- number of children (12–14 years)	0.003	0.036	0.085**	-0.028	0.006
Sex and age groups (references: 25–34 years):					
- 15–24 years	-0.061***	-0.023	0.255	0.011	-0.029*
- 35–44 years	0.023	-0.078***	0.058	-0.034	0.006
- 45–54 years	0.068***	-0.139***	0.028	0.034	-0.013
- 55–64 years	0.096***	-0.130***	0.085	0.007	-0.022*
- 65–74 years	-0.121	-0.162***		0.131	0.012
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	-0.100	-0.003	0.194	-0.113	-0.028*
- Humanities, languages, arts	0.036	0.178*	-0.008	0.034	0.032
- Science, mathematics, computing	0.314***	-0.061	-0.057	0.055	0.003
- Engineering, manufacturing, construction	0.027	0.088**	-0.064*	0.032	0.085**
- Agriculture, veterinary	-0.043	0.180**	0.013	-0.042	0.131
- Health, welfare	0.135**	-0.028	0.189	-0.055	-0.033***
- Services	-0.048**	0.070	-0.025	0.037	0.010
Observations	48,709	9,704	2,992	4,029	10,523

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Table A 12: Estimation results medium-skilled workers: Portugal, Romania, Sweden, Slovakia, and the United Kingdom

	Probability of overeducation: marginal effects				
	Medium-skilled workers aged 15–74				
	<i>PT</i>	<i>RO</i>	<i>SE</i>	<i>SK</i>	<i>UK</i>
Personal characteristics					
Female	0.034	-0.038***	0.056***	-0.025***	-0.015
Age groups (reference: 25–34 years):					
- 15–24 years	0.384***	-0.016**	-0.039***	-0.002	0.007
- 35–44 years	-0.307***	-0.004	-0.005	0.001	0.027
- 45–54 years	-0.407***	-0.000	-0.013***	0.002	0.018
- 55–64 years	-0.361***	0.017**	-0.005	0.008**	0.050*
- 65–74 years	-0.468***	0.064	-0.043***	0.008	0.054
Married	0.152	0.010**	-0.006*	-0.008***	-0.019
Foreigner: EU countries	0.405***		0.035***	0.213*	0.250***
Foreigner: non-EU countries	0.438***		0.161***		0.078
Field of study (reference: social sciences, business and law):					
- Teaching, education	0.069	-0.013	0.242***	0.300**	0.062
- Humanities, languages, arts	0.358***	-0.041***	0.037***	0.007	0.056
- Science, mathematics, computing	0.290**	-0.042***	0.156***	-0.003	0.035
- Engineering, manufacturing, construction	0.081	-0.100***	-0.020***	-0.018***	-0.033*
- Agriculture, veterinary	0.180	-0.048***	-0.025***	-0.007***	-0.073**
- Health, welfare	-0.528***	0.039	0.028***	0.078**	-0.100***
- Services	0.280**	-0.026***	-0.040***	-0.006***	-0.155***
Household characteristics					
Number of unemployed adults ¹	0.042	0.002		-0.000	0.022
Number of inactive persons ¹	-0.051	-0.003		0.002**	0.006
Persons aged 75 or older ¹					0.214
Number of children between 0 and 5 years ¹	-0.120	-0.008		-0.001	0.000
Number of children between 6 and 11 years ¹	-0.111	0.005		0.001	0.012
Number of children between 12 and 14 years ¹	-0.020	0.001		0.004**	0.002
Job characteristics					
Firm size (reference: < 10 persons):					
- 11–19 persons	-0.205**	0.003	-0.006	0.002	0.026
- 20–49 persons	0.057	-0.001	-0.013***	0.008***	0.020
- 50 and more persons	-0.081	0.011**	-0.004	0.003	0.037***
Marginal employment ²	-0.462***		-0.013	-0.004**	-0.013
Temporary contract	0.066	0.001	0.036***	0.031***	0.059**
Usual working hours	-0.363**	-0.025	-0.015	-0.027***	-0.032*
Usual working hours squared	0.033**	0.002	0.000	0.003***	0.002
Tenure	0.021	-0.008	-0.019***	-0.002	-0.034**
Tenure squared	-0.004	0.004***	-0.001	-0.000	0.007

Participation in LLL	-0.178***	-0.004	0.003	0.003	-0.023*
Second job	-0.178	-0.007	0.004	0.013	0.064**
Degree of urbanisation	0.007	-0.018***	-0.010***	0.001	0.019**
Interaction terms:					
Sex and married	-0.193	-0.013*	0.000	0.010***	0.033
Sex and urbanisation	-0.068	0.012***	-0.006**	0.003**	0.005
Sex and elder household members					
Sex and children:					
- number of children (0–5 years)	0.184*	0.019**		0.001	-0.023
- number of children (6–11 years)	0.185*	-0.002		-0.001	-0.013
- number of children (12–14 years)	-0.122	-0.003		-0.005*	0.038
Sex and age groups (references: 25–34 years):					
- 15–24 years	0.192	0.006	-0.008	-0.007***	-0.066***
- 35–44 years	0.254*	-0.010	-0.028***	-0.006***	-0.014
- 45–54 years	0.364***	-0.017**	-0.035***	-0.005***	0.007
- 55–64 years	0.276*	-0.023***	-0.039***	-0.009***	-0.042
- 65–74 years			-0.051***	-0.008***	-0.060
Sex and field of study (reference: social sciences, business and law):					
- Teaching, education	-0.212	-0.044***	-0.041***	-0.009***	-0.026
- Humanities, languages, arts	-0.035	0.013	-0.039***	0.031	-0.022
- Science, mathematics, computing	-0.014	0.013	-0.023**	0.006	-0.057
- Engineering, manufacturing, construction	-0.044	0.006	0.054***	0.020**	0.024
- Agriculture, veterinary	0.026	0.041	-0.029***	0.091**	0.021
- Health, welfare	0.490***	0.038	-0.046***	-0.002	0.033
- Services	-0.012	0.039*	0.013	0.024**	-0.010
Observations	913	29,606	63,449	20,274	7,758

¹: in same household

²: usually working less than 15 hours per week

Notes: Malta, Poland and Slovenia are excluded due to data restrictions. *, **, *** statistically significant at the 10 per cent, the 5 per cent, the 1 per cent level. Robust standard errors. Dummies for industry included. LLL: Life long learning.

Sources: EU-LFS (2013), HWWI (2015).

Appendix – Module 2

Table A 13: Composition of the explained gender pay gaps at country level (in %), 2010

	Belgium	Bulgaria	Czech Rep.	Germany	Estonia
Hours of work	3.28	0.12	0.32	5.04	1.53
Public control (>50%)	-0.61	0.06	-0.95	-0.26	0.32
Temporary contract	0.14	-0.01	0.24	0.17	0.11
Firm size	-0.31	1.68	0.03	-0.28	-2.57
Age	0.26	-0.15	-0.10	0.06	1.11
Tenure	0.24	-1.21	0.20	0.05	-0.87
Education	0.03	-1.92	0.38	1.63	-1.04
Occupation	-4.38	-1.76	-0.26	-0.55	4.32
Industry	4.09	4.49	3.55	8.64	7.28
<i>Total explained gap</i>	<i>2.76</i>	<i>1.31</i>	<i>3.41</i>	<i>14.5</i>	<i>10.19</i>
	Spain	Finland	France	Greece	Croatia
Hours of work	0.14	-0.01	-0.93	0.5	-0.14
Public control (>50%)	-0.39	2.68	-1.02	-0.5	-1.38
Temporary contract	0.00	0.47	0.25	0.13	0.31
Firm size	-1.14	-0.74	-0.09	0.53	-0.3
Age	0.68	-0.45	0.17	1.65	-0.19
Tenure	1.32	0.45	0.06	0.79	0.05
Education	-1.72	-0.22	-0.57	-1.49	-2.84
Occupation	3.58	3.38	1.29	1.08	-2.5
Industry	2.93	3.8	5.66	2.77	1.01
<i>Total explained gap</i>	<i>5.39</i>	<i>9.37</i>	<i>4.82</i>	<i>5.47</i>	<i>-5.98</i>
	Hungary	Italy	Lithuania	Latvia	Netherlands
Hours of work	-1.4	2.45	0.68	0.42	2.94
Public control (>50%)	-0.27	-0.78	-1.56	-0.75	2.39
Temporary contract	0.02	0.16	0.00	0.04	-0.08
Firm size	-2.08	-1.00	-1.7	-4.32	-0.30
Age	-0.04	-0.03	0.01	0.45	0.84
Tenure	-0.57	-0.17	-3.1	-1.48	0.40
Education	-2.44	-3.03	-2.29	-2.63	-0.53
Occupation	0.96	-9.53	-4.66	-4.63	1.53
Industry	6.26	5.69	4.55	9.71	0.00
<i>Total explained gap</i>	<i>0.45</i>	<i>-6.24</i>	<i>-8.07</i>	<i>-3.19</i>	<i>7.2</i>

	Norway	Poland	Portugal	Romania	Sweden
Hours of work	2.01	0.08	-0.2	0.06	0.41
Public control (>50%)	0.65	-1.1	-1.78	-1.96	1.97
Temporary contract	0.02	-0.25	-0.22	0.01	0.00
Firm size	-0.74	1.44	-0.94	-0.70	-0.62
Age	0.45	-0.54	-0.12	0.11	-0.23
Tenure	0.22	-0.54	-0.01	-0.78	-0.27
Education	-0.18	-3.72	-5.45	-2.19	-1.47
Occupation	0.38	-8.32	-0.17	-5.03	-0.82
Industry	4.74	5.11	7.95	11.32	7.33
<i>Total explained gap</i>	<i>7.55</i>	<i>-7.84</i>	<i>-0.93</i>	<i>0.84</i>	<i>6.29</i>

	Slovak Rep.	United Kingdom
Hours of work	0.38	0.67
Public control (>50%)	0.34	-1.25
Temporary contract	0.00	0.19
Firm size	-0.2	-0.65
Age	-0.39	-0.10
Tenure	-0.24	0.21
Education	-0.61	-0.12
Occupation	-2.69	3.91
Industry	5.60	3.10
<i>Total explained gap</i>	<i>2.19</i>	<i>5.97</i>

Sources: SES (2010), HWWI (2015).

Table A 14: Drivers of the occupation-related endowment effect (cross-country estimation) in SES (2010)

Classifications	Occupational groups	Effect (in %)
ISCO 23	Teaching professionals	-3.67
ISCO 22	Health professionals	-2.32
ISCO 32	Health associate professionals	-1.33
ISCO 33	Business and administration associate professionals	-1.05
ISCO 41	General and keyboard clerks	-0.83
ISCO 53	Personal care workers	-0.81
ISCO 52	Sales workers	-0.60
ISCO 26	Legal, social and cultural professionals	-0.36
ISCO 42	Customer services clerks	-0.34
ISCO 34	Legal, social, cultural and related associate professionals	-0.23
ISCO 51	Personal service workers	-0.21
ISCO 24	Business and administration professionals	-0.10
ISCO 54	Protective services workers	-0.09
ISCO 44	Other clerical support workers	-0.08
ISCO 43	Numerical and material recording clerks	-0.06
ISCO 96	Refuse workers and other elementary workers	-0.04
ISCO 94	Food preparation assistants	-0.03
ISCO 92	Agricultural, forestry and fishery labourers	-0.01
ISCO 2	Non-commissioned armed forces officers	0.00
ISCO 3	Armed forces occupations, other ranks	0.00
ISCO 61	Market-oriented skilled agricultural workers	0.00
ISCO 62	Market-oriented skilled forestry, fishery and hunting workers	0.00
ISCO 63	Subsistence farmers, fishers, hunters and gatherers	0.00
ISCO 95	Street and related sales and service workers	0.00
ISCO 1	Commissioned armed forces officers	0.01
ISCO 75	Food processing, wood working, garment and other craft workers	0.01
ISCO 73	Handicraft and printing workers	0.06
ISCO 82	Assemblers	0.06
ISCO 14	Hospitality, retail and other services managers	0.21
ISCO 35	Information and communications technicians	0.33
ISCO 81	Stationary plant and machine operators	0.34
ISCO 91	Cleaners and helpers	0.42
ISCO 11	Chief executives, senior officials and legislators	0.47
ISCO 74	Electrical and electronic trades workers	0.51
ISCO 71	Building and related trades workers, excluding electricians	0.68
ISCO 12	Administrative and commercial managers	0.74
ISCO 83	Drivers and mobile plant operators	0.84
ISCO 25	Information and communications technology professionals	1.06
ISCO 13	Production and specialised services managers	1.08
ISCO 72	Metal, machinery and related trades workers	1.29
ISCO 21	Science and engineering professionals	1.45
ISCO 31	Science and engineering associate professionals	1.83

Sources: SES (2010), HWWI (2015).

Table A 15: Drivers of the industry-related endowment effect (cross-country estimation)

Classification	Industry	Effect (in %)
Nace 75_86_to_88	Health and social work activities	-0.42
Nace 47	Retail trade	-0.16
Nace I	Accommodation and food services	-0.07
Nace 10_to_13 + 14_15	Food industry and textiles	-0.05
Nace 68_72_to_74_77_95 + 90_to_93_96	Professional, scientific and creative services	-0.04
Nace 94	Activities of membership organisations	-0.04
Nace 70_71_78_81_82 + 64_to_66_69_80 + 53_61_to_63_79	Business services	0.06
Nace 49_to_52	Transportation and storage	0.14
Nace 16_to_18 + 58_to_60	Paper, printing and publishing	0.21
Nace 45_46	Wholesale trade	0.71
Nace B + 35_36 + 37_to_39	Mining, energy and water supply	0.79
Nace 24_25 + 28	Basic metals and metal products	1.00
Nace 26_to_27_33 + 19_to_22 + 23 + 29_30 + 31_32	Chemical products, electric and transport equipment	1.49
Nace F	Construction	1.59

Notes: The reference group is the worst paid compared to all other industries listed here. Therefore, the sign of the effects is exclusively determined by the relative employment shares of men and women.

Sources: SES (2010), HWWI (2015).

Table A 16: Composition of the unexplained gender pay gaps at country level

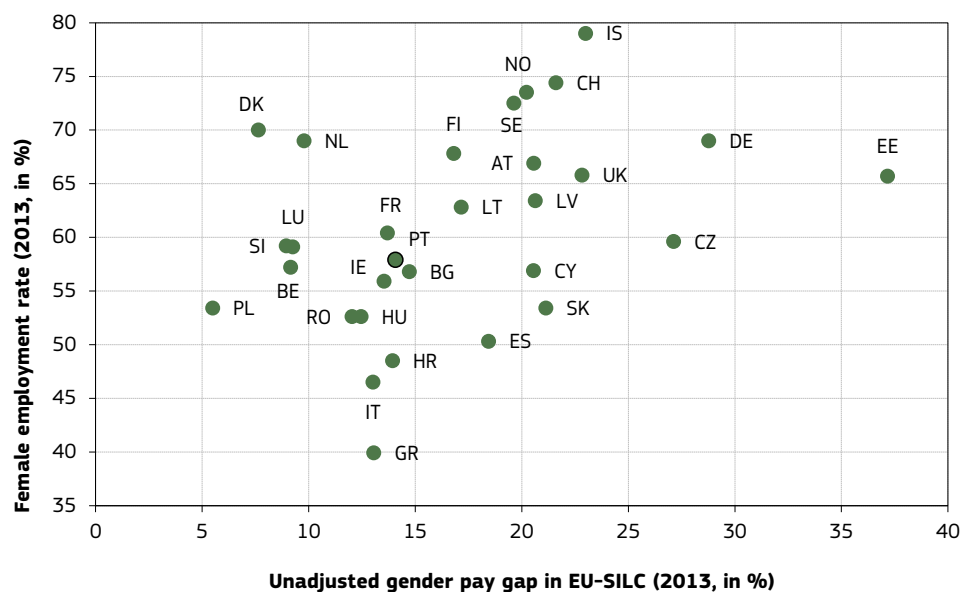
	Belgium	Bulgaria	Czech Rep.	Germany	Estonia
Hours of work	-0.95	-0.68	-0.08	-3.26	-1.55
Public control (>50%)	0.82	1.82	0.66	-0.72	3.03
Temporary contract	-0.29	0.22	0.07	-0.41	0.51
Firm size	-0.08	-1.02	-3.41	-0.20	-4.13
Age	-0.44	-1.46	-2.43	-1.70	-1.07
Tenure	-0.67	-0.76	-0.79	-0.88	0.07
Education	1.03	0.80	-0.13	-0.90	-0.05
Occupation	-2.80	9.29	1.55	4.25	-2.08
Industry	5.53	13.70	10.38	9.09	6.99
Constant	3.62	-14.59	7.31	2.47	13.16
<i>Total unexplained gap</i>	<i>5.78</i>	<i>7.32</i>	<i>13.12</i>	<i>7.74</i>	<i>14.88</i>
	Spain	Finland	France	Greece	Croatia
Hours of work	0.37	-0.09	0.80	-0.41	0.21
Public control (>50%)	-0.40	-0.88	2.85	1.07	2.19
Temporary contract	-0.31	-0.76	-0.02	-0.68	-0.23
Firm size	-2.24	-0.33	0.03	-2.18	-0.83

Age	-0.59	-0.87	-1.24	-2.75	-0.35
Tenure	-0.89	0.31	-1.62	0.96	-0.19
Education	-2.36	0.25	0.91	1.72	2.13
Occupation	-1.12	-0.25	-3.51	-2.13	-16.98
Industry	6.39	7.53	3.58	2.39	4.84
Constant	13.17	6.46	6.92	9.62	20.94
<i>Total unexplained gap</i>	<i>12.03</i>	<i>11.36</i>	<i>8.68</i>	<i>7.62</i>	<i>11.72</i>
	Hungary	Italy	Lithuania	Latvia	Netherlands
Hours of work	-1.02	-1.08	-3.31	-1.96	-3.58
Public control (>50%)	-0.17	-4.28	2.78	3.58	-2.79
Temporary contract	-0.11	-0.47	-0.13	-0.33	-0.43
Firm size	-7.96	-0.64	-5.43	-4.82	0.19
Age	-2.52	-0.45	-1.43	-0.18	-3.83
Tenure	-1.40	-0.98	1.44	-0.40	-1.26
Education	2.14	-0.07	-1.49	-1.06	1.28
Occupation	-4.26	7.02	3.68	2.54	-1.11
Industry	4.53	2.29	9.00	10.98	0.76
Constant	18.68	9.36	8.78	2.26	18.78
<i>Total unexplained gap</i>	<i>7.90</i>	<i>10.70</i>	<i>13.89</i>	<i>10.61</i>	<i>8.00</i>
	Norway	Poland	Portugal	Romania	Sweden
Hours of work	-2.30	-0.59	-0.21	-0.09	-0.13
Public control (>50%)	-2.56	-0.22	0.38	5.73	-0.91
Temporary contract	-0.10	-0.75	-0.18	0.23	0.00
Firm size	-0.32	-3.95	-1.22	-0.24	-0.21
Age	-1.29	-0.09	-2.33	-0.36	-1.67
Tenure	0.31	0.49	-1.20	-1.29	-0.04
Education	-0.69	-0.41	-0.73	0.41	0.78
Occupation	2.73	-2.97	-2.44	3.15	-0.36
Industry	3.62	3.84	4.32	11.77	0.01
Constant	7.35	16.05	15.94	-13.10	10.25
<i>Total unexplained gap</i>	<i>6.74</i>	<i>11.40</i>	<i>12.32</i>	<i>6.21</i>	<i>7.72</i>
	Slovak Rep.	United Kingdom			
Hours of work	-0.24	2.07			
Public control (>50%)	-0.61	-3.03			
Temporary contract	0.80	-0.78			
Firm size	-2.17	-1.63			
Age	-2.46	-6.17			
Tenure	-1.70	-0.73			
Education	-0.40	0.76			
Occupation	3.89	-0.89			
Industry	4.01	-1.64			
Constant	13.31	26.31			
<i>Total unexplained gap</i>	<i>14.43</i>	<i>14.29</i>			

Sources: SES (2010), HWWI (2015).

Appendix – Module 3

Figure A 1: Relationship between gender pay gap and female employment in EU-SILC



Sources: Eurostat (2015), EU-SILC (2013), HWWI (2015).

Table A 17: Unadjusted, explained and unexplained gender pay gap in EU-SILC (in %), 2013

Country	Unadjusted gap	Explained gap	Unexplained gap (adj.)	Country	Unadjusted gap	Explained gap	Unexplained gap (adj.)
Austria	20.56	15.61	4.95	Latvia	20.64	3.25	17.39
Belgium	9.16	-2.78	11.94	Lithuania	17.16	-5.40	22.56
Bulgaria	14.73	-0.62	15.34	Luxembourg	9.26	-6.73	15.99
Croatia	13.94	-5.04	18.98	Netherlands	9.78	-8.56	18.35
Cyprus	20.56	12.57	7.99	Norway	20.22	16.27	3.96
Czech Rep.	27.13	4.95	22.18	Poland	5.50	-3.31	8.80
Denmark	7.64	-2.41	10.05	Portugal	14.07	0.26	13.81
Estonia	37.16	12.88	24.29	Romania	12.04	-1.46	13.49
Finland	16.81	6.21	10.60	Serbia	-0.04	-4.41	4.37
France	13.70	14.68	-0.98	Slovakia	21.14	1.44	19.70
Germany	28.78	19.45	9.33	Slovenia	8.96	-9.74	18.69
Greece	13.06	4.31	8.75	Spain	18.45	0.90	17.56
Hungary	12.47	-0.38	12.84	Sweden	19.63	6.21	13.42
Iceland	23.00	10.74	12.26	Switzerland	21.61	20.95	0.65
Ireland	13.54	-1.07	14.61	UK	22.83	15.31	7.52
Italy	13.02	3.75	9.27	Total	17.99	6.85	11.14

Sources: EU-SILC (2013), HWWI (2015).

Table A 18: Composition of the explained gender pay gaps at country level in % (2013)

	Austria	Belgium	Bulgaria	Croatia	Cyprus
Hours of work	3.79	-7.11	-0.24	1.17	1.15
Age	-1.37	-0.34	-0.14	-0.18	1.84
Education	-0.90	-0.92	-6.68	-2.58	-0.40
Temporary contract	0.01	0.12	0.00	0.81	1.10
Health	-0.09	0.14	0.09	-0.33	0.00
Overeducation	0.30	0.03	0.18	0.06	-0.05
Supervisor	3.03	1.03	0.12	0.38	2.67
Partner context	-0.48	-1.51	-4.15	-1.73	-0.84
Size of the local unit	2.64	0.89	1.28	1.64	-0.67
Country of birth	0.12	0.09	-0.07	-0.10	0.91
Occupation	3.76	-0.25	6.00	-8.17	10.81
Industry	4.82	5.05	2.99	3.99	-3.95
<i>Total explained gap</i>	<i>15.61</i>	<i>-2.78</i>	<i>-0.62</i>	<i>-5.04</i>	<i>12.57</i>
	Czech Rep.	Denmark	Estonia	Finland	France
Hours of work	0.23	1.96	1.85	0.53	7.72
Age	-0.09	-0.48	1.55	-0.75	-0.84
Education	0.50	-2.33	-3.48	-1.63	-1.14
Temporary contract	0.45	-0.27	0.08	2.32	0.44
Health	0.08	-0.05	0.26	-0.15	0.27
Overeducation	0.01	-0.05	0.75	0.14	0.03
Supervisor	1.35	0.29	0.69	0.98	0.47
Partner context	-2.19	-0.11	-0.03	-0.34	0.52
Size of the local unit	1.05	-0.07	-0.22	0.53	1.38
Country of birth	0.01	-0.05	0.05	0.00	-0.26
Occupation	0.86	3.20	5.45	0.43	0.28
Industry	2.68	-4.46	5.94	4.14	5.80
<i>Total explained gap</i>	<i>4.95</i>	<i>-2.41</i>	<i>12.88</i>	<i>6.21</i>	<i>14.68</i>
	Germany	Greece	Hungary	Iceland	Ireland
Hours of work	5.07	0.28	-0.36	-0.10	-5.26
Age	-0.84	2.30	-0.58	-0.33	0.00
Education	1.28	-2.15	-4.35	-3.69	-0.52
Temporary contract	0.32	0.08	0.07	1.02	0.00
Health	0.20	-0.27	0.35	-0.23	0.08
Overeducation	0.56	0.43	0.40	-0.05	0.17
Supervisor	0.92	1.00	0.92	1.15	1.07
Partner context	0.00	-0.20	-0.06	0.62	2.09
Size of the local unit	5.22	0.92	0.46	-2.26	2.22
Country of birth	-0.01	0.09	0.03	0.14	0.04
Occupation	-0.72	1.30	1.48	6.80	-3.62
Industry	7.45	0.53	1.28	7.67	2.66
<i>Total explained gap</i>	<i>19.45</i>	<i>4.31</i>	<i>-0.38</i>	<i>10.74</i>	<i>-1.07</i>

	Italy	Latvia	Lithuania	Luxembourg	Netherlands
Hours of work	0.70	-0.29	1.64	-12.22	-18.66
Age	-0.14	-0.06	-1.54	0.62	-0.40
Education	-2.28	-6.21	-2.01	-0.85	-0.42
Temporary contract	0.55	-0.28	-0.45	0.14	0.13
Health	0.13	1.04	0.07	0.21	0.52
Overeducation	-0.20	0.77	0.02	-0.05	-0.14
Supervisor	1.55	0.21	0.53	2.02	1.87
Partner context	-1.79	1.37	-2.48	-1.14	-1.46
Size of the local unit	1.91	1.45	-0.27	2.69	0.36
Country of birth	0.29	-0.01	0.40	-0.26	-0.28
Occupation	-0.02	-0.59	-7.28	0.18	4.86
Industry	3.05	5.87	5.96	1.95	5.05
<i>Total explained gap</i>	<i>3.75</i>	<i>3.25</i>	<i>-5.40</i>	<i>-6.73</i>	<i>-8.56</i>
	Norway	Poland	Portugal	Romania	Serbia
Hours of work	-2.11	0.07	-1.44	-0.24	-0.39
Age	-0.30	0.09	-0.09	0.39	0.17
Education	-1.11	-3.42	-3.14	-2.00	-1.38
Temporary contract	2.29	0.12	0.42	-0.06	0.00
Health	0.02	0.17	0.58	0.16	0.26
Overeducation	-0.01	0.03	0.02	-0.16	0.00
Supervisor	3.40	0.39	1.07	0.36	0.95
Partner context	-1.55	-1.09	-0.98	-0.34	-1.05
Size of the local unit	0.81	1.18	-0.06	1.11	0.89
Country of birth	0.49	-0.01	-0.03	-0.03	0.00
Occupation	8.69	-4.17	-0.30	-0.94	-7.53
Industry	5.64	3.34	4.21	0.28	3.68
<i>Total explained gap</i>	<i>16.27</i>	<i>-3.31</i>	<i>0.26</i>	<i>-1.46</i>	<i>-4.41</i>
	Slovakia	Slovenia	Spain	Sweden	Switzerland
Hours of work	-0.34	-0.94	-2.10	4.27	10.66
Age	-0.11	-0.36	0.40	-1.25	0.14
Education	-0.81	-4.92	-0.68	-2.80	3.58
Temporary contract	0.17	0.00	1.16	1.50	0.08
Health	0.28	1.22	0.11	0.29	0.10
Overeducation	0.11	0.57	0.09	0.16	-0.14
Supervisor	0.63	0.58	1.11	0.65	2.09
Partner context	0.47	-1.00	-0.39	-2.97	0.30
Size of the local unit	0.96	-0.59	1.28	0.50	1.05
Country of birth	-0.05	-0.18	0.16	-0.14	0.19
Occupation	-3.16	-6.07	-1.50	-0.92	0.29
Industry	3.29	1.97	1.25	6.91	2.59
<i>Total explained gap</i>	<i>1.44</i>	<i>-9.74</i>	<i>0.90</i>	<i>6.21</i>	<i>20.95</i>

	United Kingdom
Hours of work	2.30
Age	-0.41
Education	0.28
Temporary contract	0.08
Health	0.00
Overeducation	-0.03
Supervisor	1.62
Partner context	0.07
Size of the local unit	0.82
Country of birth	0.01
Occupation	2.95
Industry	7.63
<i>Total explained gap</i>	<i>15.31</i>

Sources: EU-SILC (2013), HWWI (2015).

Table A 19: Drivers of the occupation-related endowment effect (cross-country estimation) in EU-SILC

<i>Classifications</i>	<i>Occupational groups</i>	<i>Effect (in %)</i>
ISCO 23	Teaching professionals	-2.59
ISCO 22	Health professionals	-1.55
ISCO 41	General and keyboard clerks	-1.14
ISCO 32	Health associate professionals	-1.03
ISCO 33	Business and administration associate professionals	-0.84
ISCO 52	Sales workers	-0.69
ISCO 42	Customer services clerks	-0.44
ISCO 53	Personal care workers	-0.44
ISCO 91	Cleaners and helpers	-0.25
ISCO 26	Legal, social and cultural professionals	-0.24
ISCO 34	Legal, social, cultural and related associate professionals	-0.16
ISCO 51	Personal service workers	-0.14
ISCO 24	Business and administration professionals	-0.07
ISCO 44	Other clerical support workers	-0.06
ISCO 54	Protective services workers	-0.05
ISCO 94	Food preparation assistants	-0.05
ISCO 92	Agricultural, forestry and fishery labourers	-0.01
ISCO 61	Market-oriented skilled agricultural workers	-0.01
ISCO 95	Street and related sales and service workers	0
ISCO 43	Numerical and material recording clerks	0
ISCO 96	Refuse workers and other elementary workers	0
ISCO 63	Subsistence farmers, fishers, hunters and gatherers	0
ISCO 2	Non-commissioned armed forces officers	0
ISCO 3	Armed forces occupations, other ranks	0
ISCO 1	Commissioned armed forces officers	0
ISCO 62	Market-oriented skilled forestry, fishery and hunting workers	0
ISCO 75	Food processing, wood working and garment	0
ISCO 82	Assemblers	0.07
ISCO 73	Handicraft and printing workers	0.09
ISCO 14	Hospitality, retail and other services managers	0.11
ISCO 81	Stationary plant and machine operators	0.11
ISCO 35	Information and communications technicians	0.19
ISCO 12	Administrative and commercial managers	0.24
ISCO 11	Chief executives, senior officials and legislators	0.31
ISCO 74	Electrical and electronic trades workers	0.51
ISCO 13	Production and specialised services managers	0.54
ISCO 83	Drivers and mobile plant operators	0.71
ISCO 71	Building and related trades workers, excluding electricians	0.74
ISCO 25	Information and communications technology professionals	0.96
ISCO 72	Metal, machinery and related trades workers	1.23
ISCO 21	Science and engineering professionals	1.31
ISCO 31	Science and engineering associate professionals	1.52

Note: Reference group: Labourers in mining, construction, manufacturing and transport.

Sources: EU-SILC (2013), HWWI (2015).

Table A 20: Drivers of the industry-related endowment effect (cross-country estimation) in EU-SILC

Classification	Industry	Effect (in %)
G	Wholesale and retail trade	-0.35
K	Financial and insurance activities	-0.28
R-U	Other service activities	-0.20
L-N	Professional and scientific activities	-0.19
A	Agriculture, forestry and fishing	-0.01
I	Accommodation and food service activities	0.08
Q	Human health and social work activities	0.48
J	Information and communication	0.53
H	Transporting and storage	0.85
F	Construction	1.13
B-E	Mining, manufacturing and energy	2.99

Note: Reference group: Education. Sources: EU-SILC (2013), HWWI (2015). Tis doluptassit, ullores sit, int.

