JRC TECHNICAL REPORTS

Over-education of migrants? Evidence from the EU

Biagi, F.; Grubanov, S. and Mazza, J.

2019

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

Contact information
Name: Jacopo MAZZA
Address: Via E. Fermi, 2749, 21027 Ispra VA
Email: Jacopo.mazza@ec.europa.eu
Tel.:+39 0332 78 6318

Contents

Authors

Biagi Federico[*], Grubanov-Boskovic Sara[+], Jacopo Mazza[+]

[+]European Commission, Joint Research Centre, Knowledge Centre on Migration and Demography
[*] European Commission, Joint Research Centre

Abstract

Using five waves of the European Labour Force Survey for the period 2011-2016 we analyse the differential incidence of overeducation between natives and migrants in twenty-four EU Member States. We look separately at secondary and tertiary educated individuals and for the latter group we apply two separate methods to measure overeducation: the Eurostat method and the realized matches method. We also look at how the likelihood of being over/undereducated is influenced by the length of stay in the host country.

In the first part of our analysis we present simple descriptive statistics, and document that, on average, non-EU born (NEB) and European migrants are less well matched than natives with comparable (i.e. secondary and tertiary) education. However, these basic descriptive statistics are likely to be affected by how individual characteristics (potentially affecting overeducation) are distributed among native and migrant workers in the sample used in the analysis. Hence, in the second part, we exploit the rich set of information provided by the LFS and probe our data further. Applying standard econometric techniques, which allow us to control for observable characteristics (year, country and industry fixed effects, age and its squared term, the degree of urbanization in the area of workers' residence and gender), we test whether being a migrant, *per se*, can significantly affect the likelihood of being over/undereducated. Our results confirm that EU migrants and NEBs (*ceteris paribus*) are more likely to be overeducated (and less likely to be undereducated) compared to natives with the same educational level.

Our data also allows us to check whether the quality of the match improves or degrades as time of residence in the host country increases. We find that the negative gap for both secondary and tertiary educated NEBs and EU migrants increases with the length of stay in the host country with the exception of EU migrants with short tertiary attainment, for which the opposite holds.

Several not mutually exclusive interpretations can be advanced for our results, for example: even if equipped with the same amount of formal education, NEB migrants might lack in other dimensions of their human capital; their social network might be less extensive than those of locals negatively affecting their chance of finding a good match on the labour market; some discrimination on the part of local employers might be at play. Each possible explanation calls for a different mix of public policies, but our analysis does not allow for a casual interpretation of the findings which goes beyond the scope of this report.

Nonetheless, the results presented here are important for at least two reasons: first, they document how the measurement of overeducation is to some extent affected by the methodology adopted; second, they suggest that overeducation among migrants, especially the non-European ones, is persistent and does not seem to disappear with their permanence in the host country.

1 Introduction

This report analyses the prevalence of vertical skill mismatches in the EU, comparing migrants (from the EU and from outside the EU) to EU natives in the years 2011-2016. This is a very relevant aspect for many EU economic and social policies as stressed in the 2015 European Agenda on Migration (EAM), that –directly or indirectly- aim at migrants' integration in EU society and labour markets.

Vertical skill mismatches captures situations in which the skills implied by job requirements and the skills possessed by the worker can be compared on an ordinal scale (i.e., whether a worker has more or less skills than those required to perform her/his job tasks). Vertical skills mismatches are often associated to over/under education, given the relationship between skills and educational attainment levels.

While the relationship between skills and education is very complex and should not be assumed lightly, we concentrate on over/undereducation because it is very difficult to obtain reliable data on migrants' skills across the EU, while it is easy to have them on their education levels.

Overeducation can be a structural phenomenon, determined by the interaction between demographics, educational choices and technological progress (Verhaest et al., 2017). Or it could be temporary, reflecting the economic cycle and labour demand/supply shocks, as well as - in the case of migrants - specific national migration policies (OECD 2014).

Over/undereducated workers signal a potential misallocation of resources and a loss of productivity and talent, with potentially negative consequences on productivity, wages and job satisfaction (Kiker et al. 1997; Dolton and Vignoles, 2000; McGuiness, 2006; CEDEFOP 2010; Quintini, 2011). On the other hand, part of the literature stresses the need to be very careful when drawing implications on allocative efficiency from a simple measure of over/undereducation, given that workers' productivity and wages depend on a wide set of factors besides education. Innate and unmeasured ability, cognitive and non-cognitive skills, general and job-specific knowledge, personality traits- together with education- all determine workers' contribution to the production process. Nevertheless, since educational attainment is often a reflection of unobservable traits such as ability, cognitive skills and knowledge, it is interesting to explore the educational dimension of vertical mismatches, keeping in mind the dangers of drawing policy conclusions from such crude measures.

Existing studies document that overeducation is significant especially for migrants (Groot and Maassen van den Brink, 2000; Leuven and Oosterbeek, 2011). The literature has found that migrants tend to be more overeducated then natives (OECD/European Union, 2018). High overeducation among migrants is often linked with problems in getting recognition of their degrees, lower level of skills and knowledge compared to natives with the same educational achievement, lack of network connections or labour market experience necessary to find well-matched jobs, low skills in the language of the host country and, finally, labour market discrimination against foreigners.

In this paper we are not able to address the role of all these potential sources of mismatches as we lack the appropriate information at the individual level. However, with the available data, we can look at one interesting aspect of overeducation: the relevance of work experience and explore whether over/undereducation is a temporary or permanent phenomenon (i.e., whether the likelihood of being over/undereducated changes with labour market experience). Due to information asymmetries, young workers with little or no experience in the labour market are more likely to be overeducated. This differential should disappear with time since workers have the opportunity to find a better job match (i.e., overeducation should be reduced as labour market experience/permanence in the host country increases). On the other hand, time spent in an improperly matched job might depreciate workers' human capital and introduce signalling effects. This implies that

overeducation might become heavily persistent if a worker, especially when young, spends a lot of time in a poorly matched position[1] (Meroni and Vera Toscano, 2017).

Whether the likelihood of being over/undereducated changes with length of stay in the host country and whether there are significant differences between natives and migrants in this evolution are interesting questions that we address in this report. In particular, we focus on over/undereducation among two groups: i) those who have completed tertiary education (ISCED level 5 to 8); ii) those who have competed secondary or post- secondary education (ISCED level 3 and 4), paying attention to observable demographic differences between the two groups. We think that is important to look at these two groups separately as different socio-economic forces could be at play. Migrants with tertiary degrees might come from countries for which difficulties in recognition of previous studies and discrimination play a minor role and they might arrive in the host country with good job prospects, hence being positively selected into the labour market. On the other hand, for secondary education, cream-skimming might be less relevant and improper functioning of labour markets, lack of occupational-specific skills and language skills, difficulty in mutual recognition and discrimination might be more likely to exist. Whether or not and for whom these effects are at play is a matter of empirical analysis that would deserve further research. In this report we simply look at difference in over/undereducation between short and long-term residents.

We adopt a statistical approach to over/undereducation: for all education groups we use the realized matches method, comparing the actual level of education of each worker to that prevalent (i.e., the mode) among those that are employed in the same two-digit ISCO08 occupation. The mode is time and country (besides 2 digit ISCO08) specific: this seems reasonable as different countries might experience different levels of technological progress and different evolutions of the demographic and educational composition reducing, in this way, the chances of upward bias in our estimates of overeducation.

For tertiary graduates we complement the realized matches method with one developed by Eurostat (in its experimental statistics), which defines as overqualified all those workers with tertiary education who are working in occupations for which such education level is not required (i.e., those with educational attainment ISCED11 5 to 8 and employed in one-digit ISCO08 occupations 4 to 9). By computing both measures of over-education, our paper sheds some light on the drivers of the different estimates obtained using the Eurostat and the realized matches methods for tertiary graduates. The report proceeds as follows: Section 2 discusses the issue of vertical mismatches and the possible approaches to its measurement, explaining in details the ones adopted in this report. Section 3 presents the data and summary statistics, also focusing on the difference between the realized matches and the Eurostat method (for tertiary graduates). Section 4 contains our empirical analysis, which uses regression methods to estimate the role of area of origin and length of stay on the likelihood of being over/undereducated as well as being well matched. Section 5 concludes.

---

[1] This would happen if the individual who is overeducated at the beginning of the working career ends up with a lower "usable" level of human capital, relative to the one that would be normally associated to her level of education. In this case she might not be considered overskilled even if she is overeducated.

2 How to measure mismatches

Skills mismatches exist when a given worker's set of skills does not match those required by the job that she/he is performing[2]. It is useful to distinguish between vertical and horizontal skill mismatches. Vertical skill mismatches refer to a situation in which there is an ordinal relationship between the skills implied by job requirements and the skills possessed by the worker. In this case, it is appropriate to inquire whether a worker has more/less/adequate skills compared to those required by her/his job (e.g., we have vertical mismatch when a civil engineer performs the tasks of a construction worker). Horizontal skills mismatch refers to a situation in which a worker possesses skills that are different from those that are necessary to perform well her/his job tasks, but no clear ordinal relationship emerges (e.g., an electric engineer performing the tasks of an accountant).

The concept of vertical mismatch is typically classified as over or undereducation, since skills are often measured using formal educational attainment, which can easily accommodate an ordinal structure (e.g., tertiary education is "more" that "secondary education" which is "more" than primary education etc.).

However, skills and education are different concepts. Skills are defined by the OECD as "*bundle of knowledge, attributes and capacities that can be learned and that enable an individual to successfully and consistently perform an activity or task, whether broadly or narrowly conceived, and can be built upon and extended through learning*" (OECD, 2012). While this definition is not undisputed([3]), everyone agrees that skills can be acquired through formal and informal education, training and job experience/practice. Hence, formal education is just one of the components of skills acquisition, and its relevance might decrease as workers age, leaving space to training and work experience.

In fact, part of the literature has abandoned the use of formal education as a proxy for skills and has started to use surveys in which information on the relationship between the tasks performed and the ability of workers to perform them are specifically addressed ([4]). This, for instance, is the case of PIAAC, whose first wave has been used extensively to measure skill mismatch, over/underskilling, in parallel to over/undereducation ([5]).

Given our focus on migrants and our interest in EU cross-country comparisons, in this work we do not use PIAAC or other task-based surveys, as it would not be possible to have both data for all EU MS and sufficient observations to estimate differences between natives and migrants (especially if we want to distinguish between EU and non-EU migrants). We rely on Labour Force Survey data, but this has immediate implications for our measurement: we can only look at over/undereducation as opposed to over/underskilling.

There are different ways in which over/undereducation can be assessed. Some approaches are subjective, as they rely on workers' self-assessments captured through surveys, while others are objective in nature, as they use measures that do not rely of workers' perceptions.

Two are the prevailing objective approaches. The first one - typically referred to as the *Normative* or *Job Analyst* method - uses the educational requirements of occupations specified by professional job analysts to derive implication on vertical mismatches of individual workers (i.e., a worker is over/undereducated if she/he has a level of educational attainment that is higher/lower than the one specified by the job analyst for her/his occupational category). This method works well when the job descriptions, job skills requirements and the corresponding mapping into educational requirements are defined for detailed and specific occupations, are updated frequently and reflect country

---

([2]) For a recent contribution to the debate on skill mismatches, see McGuinness (2018).
([3]) Others argue that skills capture technical capacities, while knowledge, attitudes and abilities –together with skills- make up competencies, which is what labour market remunerates.
([4]) This is the case of surveys that explicitly ask workers whether they have the appropriate skills required by their tasks or whether they think that they are over/under-skilled.
([5]) See, for example Flisi et al., 2017.

specificities (Rumerger, 1987; McGoldrick and Robst, 1996). This is still not the case in the EU where such EU-wide mapping has not been formalized yet, although progress is expected with the European Skills, Competences, Qualifications and Occupations (ESCO) classification[6].

The second one, called *realized matches method*, is a statistical approach that uses information on the distribution of educational attainment within an occupation to infer over/undereducation (Verdugo and Verdugo, 1989; Kiker et al., 1997). In this case, the educational level of each worker is compared to the one prevailing (i.e., the median or the mode) among workers in the same occupation, and a worker is defined as over/undereducated if she/he has an educational attainment higher/lower than the prevailing one (often a measure of distance from the median/mode is also used). Hence, the realized matches method reflects the statistical properties of the underlying database, the choice of the reference educational level and the conditioning variables used to define the latter (i.e., occupation[7], age, gender, migrant status etc.). For instance, within the same occupation, educational attainments of workers might vary in time (and by country) due to technological change and educational reforms (e.g., reforms favouring tertiary education). It is hence frequent that, within the same occupation, young workers have higher educational attainment levels than older cohorts. This implies that young workers are more likely to appear as overeducated in occupations in which older workers tend to be numerous[8]. These peculiarities make across-country comparability more complex. A major drawback of the realized matches method is the potential endogeneity of the indicator: occupations in which overeducation would be prevalent under the Job-analyst method, would not appear so under realized matches, since the median/mode educational attainment would be high (Verhaest and Omey, 2006; CEDEFOP, 2010).

An alternative to the methods based on objective measures is to rely on subjective evaluations. In this case the analyst relies on information provided -directly or indirectly- by workers (self-assessment, self-declared etc.) on their level of education and its appropriateness in relationship to job tasks (Desjardins and Rubenson, 2011; Pellizzari and Fichen, 2013; Pellizzari et al. 2015; Flisi et al. 2017). While appealing for its attention to the specificities of individual workers, subjective approaches suffer from measurement error (workers have different interpretation of the appropriate levels of education) and from response bias (workers might tent to overstate the educational requirements of their job, in order to gain social status or simply equate job requirements to their own level of education). Additionally, workers' surveys of subjective measurement are infrequent[9]. A recent study that uses the subjective approach with data from a multi-country web survey is Visintin et al. (2015), who document that migrants in Europe and Asia are more likely to be overeducated than natives, while the opposite is true for migrants in Africa and Latin America (females are also more likely to be overeducated while the opposite holds for individuals with higher education).

Finally, the level of aggregation at which the analysis is conducted also plays an important role: the higher is the level of aggregation for occupations (i.e., ISCO one-digit) and education (e.g., secondary vs. tertiary) and the more we overlook the underlying differences, allowing for compositional effects to drive the results in a significant way.

In our work we follow the realized matches method (subjective approaches are not possible with the LFS) and – for tertiary graduates- we complement it with a method proposed by Eurostat in its experimental statistics[10], which can be interpreted as a rough approximation to the Job analyst method. Eurostat defines as overqualified workers with tertiary education who are working in occupations for which such education level is not required (i.e., those with educational attainment ISCED11 5 to 8 and employed in one-

---

[6] In the US this is possible with the use of the Occupational Information Network (O*Net).
[7] Occupations can be defined at the one, two, three or four digit levels, and this will have consequences on the reference educational attainment level.
[8] A possible response would be to compute the reference educational level distinguishing by age.
[9] An exception in the European Working Condition survey, which, however, suffers from small sample size, which is particularly relevant when performing a cross-country analysis of migrants.
[10] See https://ec.europa.eu/eurostat/web/skills/background/experimental-statistics .

digit ISCO08 occupations 4 to 9)([11]). In essence, the Eurostat definition assumes that the "normal" educational attainment in ISCO08 1-3 is tertiary education (i.e., these are "high skills" occupations), while it is secondary for ISCED08 4-8 (i.e., "medium skills" occupations) and primary for ISCO08 9 (i.e., "low skills" occupations). This is clearly a simplification, and it overlooks within one-digit ISCO08 variability. We hence show how the results obtained with the realized matches method differ from those obtained with the Eurostat method and we also clarify the origins of such differences. Our work extends European Commission (2015) (in particular its Section 4) in a significant way as we: i) use more recent data; ii) consider workers with secondary education as well as tertiary graduates; iii) complement summary statistics with an econometric analysis directed at exploring the relevance of country of birth on the likelihood of being over/undereducated or well matched; iv) discuss in depth the differences between the realized matches method results and those obtained using the Eurostat method.

([11]) One-digit ISCO08 occupations are: managers (1), professionals (2), technicians and associate professionals (3), clerical support workers (4), services and sales workers (5), skilled agricultural, forestry and fishery workers (6), craft and related trade workers (7), plant and machine operators and assemblers (8), elementary occupations (9).

3 Descriptive Statistics

For our analysis we use the European Labour Force Survey (EU-LFS). The EU-LFS is a large survey conducted quarterly in each Member State by the national institute of statistics of the Member State and centrally processed by Eurostat. It collects information on many demographic characteristics and labour force participation of people aged 15 and above.

We use data for the five-year period 2011-2016 for each member state. That gives us an initial sample of 26,921,072 observations, but we limit this initial sample in several ways. First, we keep only prime aged workers between 22 and 57 years of age. Second, we drop self-employed, armed forces and family workers. Third, we exclude respondents from Germany and Malta since the national institutes of statistics for these two countries do not collect information on whether the respondent is born inside or outside of the country, making it impossible to distinguish between migrants and natives, and from Bulgaria and Romania because of the low reliability of data on migrants([12]).

Natives are defined as those who are born in the country in which they currently work. EU migrants are EU citizens born in a EU country different from that in which they work, while NEBs are workers born outside the EU and working in a EU MS (i.e., we have adopted a definition of migrants based on the country of birth as opposed to country of citizenship([13])).

For the definition of educational level, we follow ISCED 2011([14]), but we group some levels together. We create one category including both ISCED 0 and 1 (we call it "Primary" but it includes also individuals with ISCED 0) and we group people with at least a bachelor degree into another category (ISCED 6 to 8), which we call "Tertiary". We are then left with six educational categories fairly closely replicating the ISCED 2011 classification.

After having selected the sample according to the rules specified above, we are left with 5,589,671 observations across 24 MS and 5 years.

In Figure 1 we show how our sample is distributed between countries and educational levels. In terms of geographical distribution, our sample is drawn heavily from four MSs: France, Italy, Poland and Sweden([15]). If we turn our attention to the distribution between educational levels, upper secondary is the largest group followed by tertiary. This is true for almost all MSs with the exception of: i) Ireland and Latvia where the ranking is inverted; ii) Spain where people are almost equally split between lower and upper secondary and tertiary; iii) Italy where the second largest group is that of those with lower secondary education. It is also worth noting how small is the incidence of the other three categories on the total: workers with at most completed primary education are very few everywhere, while the two categories of post-secondary and short tertiary graduates are of some relevance only in few countries (e.g., Austria, Spain, France, Greece, Ireland, Latvia, Sweden and the UK) reflecting the cross-country variation in educational systems.


([12]) See Eurostat, 2018.
([13]) Country of birth appears preferable as criteria for citizenship vary across countries.
([14]) ISCED 2011 classification has eight levels of education: 0 (Early childhood Education: 01 Early childhood educational development; 02 Pre-primary education); 1 (Primary education); 2 (Lower secondary education); 3 (Upper secondary education); 4 (Post-secondary non-tertiary education); 5 (Short-cycle tertiary education); 6 (Bachelor or equivalent); 7 (Master or equivalent); 8 (Doctoral or equivalent).
([15]) Please notice that these are unweighted observations and they are informative of the within-country distribution of educational attainments.

Figure 1. Number of individuals by educational level and Member State.
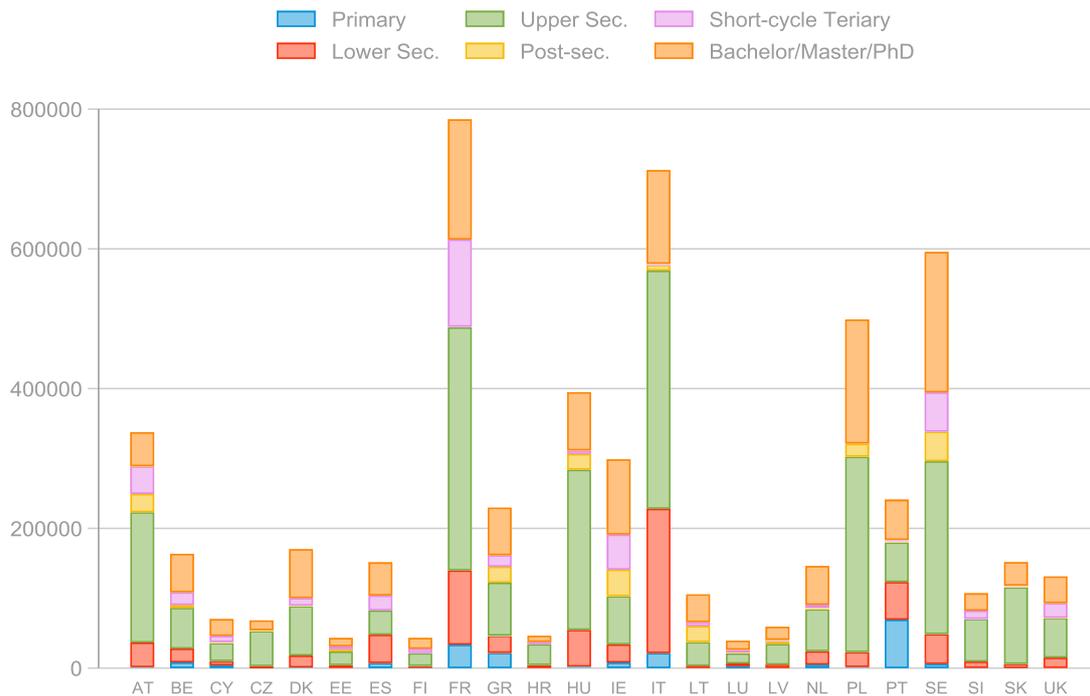
3.1.1 Over/undereducated and well matched by level of education

In this sub-section we look at over/undereducation across all 6 educational categories, and provide the full picture of over/undereducation across all the ISCO08 two-digit occupations according to the realized matches method. Following this method applied to two-digit ISCO08 occupations, a worker is identified as overeducated if he has an educational attainment level higher than the modal one (where the mode is two-digit ISCO08-country-year specific), while he is defined as undereducated in the opposite case. Otherwise, the worker is considered well matched.

An interesting picture can be obtained by looking at the shares of overeducated, well matched and undereducated workers for different levels of education and by migratory status. We can notice (Figure 2) that, among those with at most completed primary education, undereducation is widespread (and more so for natives([16])). For this group overeducation is simply not possible by default. A symmetric situation is found for tertiary graduates: in this case overeducation is widespread (and more so for NEB), while –by construction- undereducation is not possible([17]). As the educational attainment rises from lower secondary to short-cycle tertiary, we can notice that the prevalence of overeducation increases as well. Among those with lower secondary educational attainment, undereducation is prevalent over being well matched or overeducated, and especially so for natives. Among those who have completed upper secondary education, on the other hand, most workers are well matched (especially natives), but overeducation starts to

---

([16]) For each group (defined in terms of region of origin) we compute the ratio of those that are over/under-educated and well matched over those that belong to the group.

([17]) Since here we consider only ISCED levels 6 to 8, anyone who has attained these levels of tertiary education has also reached the highest possible educational level. On the other hand, we noticed that we have undereducation among ISCED level 5 holders when the modal educational attainment in their two-digit occupation is ISCED 6-8.

grow, especially among NEB and EU migrants. For those who have attained post-secondary education, overeducation is the prevailing condition, especially for NEB, and the share of well-matched in this educational category is small. Finally, for those with short-tertiary educational attainment overeducation is prevalent especially for NEB and EU migrants and the shares of well-matched reaches the lowest values for NEB and EU migrants as well. However, in this group, we also have a significant number of workers who are undereducated (they work in occupations in which ISCED 6-8 is the modal educational attainment).

Figure 2 shows that important differences exist not only between secondary and tertiary education, but also within each of these aggregates: between secondary (ISCED 2-3) and post-secondary (ISCED 4) education, and between short tertiary (ISCED 5) and tertiary (ISCED 6-8) education. This drives us to structure our empirical analysis of secondary and tertiary graduates so as to take into account those differences.

Figure 2. Realized matches by educational level and region of origin.

3.1.2 Realized matches vs. Eurostat

The purpose of this sub-section is to document the differences that exist between the realized matches method and the Eurostat method when it comes to determine the over (under) education status of workers.

Our strategy here is to compare the "objective" appropriate educational attainment according to Eurostat (for one-digit ISCO08 occupations) to those that emerge when aggregating at the one-digit level the prevailing modal educational attainments observed at the two-digit ISCO08 level.

According to Eurostat, which consider occupations at the one-digit level (i.e., from ISCO08 1 to 9), all workers with a tertiary educational attainment are well matched if they work in ISCO08 1-3 and are over-educated if they work in ISCO08 4-9[18]. So, by construction, there is no within one-digit ISCO08 level variation. On the other hand, in the realized matches approach it is possible to look within one-digit ISCO levels (i.e., at the two-digit level) and so it is interesting to compare the results obtained using the two methods([19]), and hence their robustness. In particular, we are interested in considering the extent to which the educational attainment modes observed in two-digit ISCO08 occupations (according to the realized matches approach) corresponds to the assumptions of the Eurostat method.

As expected, we find that the result of the two methods differ substantially. When we apply the realized matches approach to ISCO08 two-digit occupations and aggregate the results at the one-digit level, we get a picture quite different from that obtained using the Eurostat method (see Figure 3). Discordance between the realized matches and the Eurostat method are particularly high for ISCO08 one-digit occupations 1, 4 and –especially- 3 (in which significant shares of two-digit occupations have secondary education as their mode([20])). Some differences are also observed for one-digit ISCO08 occupations 5-9([21]).

Notice that important differences also exist between countries, reflecting different levels of educational attainment and technological progress.

Figure 4 provides a figure similar to Figure 3 but focuses on individuals as opposed to occupations. That is, it provides information of the shares of workers with different educational achievements across the one-digit ISCO08 occupations. The interest in looking at educational achievement of individuals -as opposed to the mode within two- digit occupations- come from the fact that here we are not affected by the potential endogeneity of the realized matches method([22]). In Fact, Figure 4 is a snapshot of the distribution of individual educational attainments within each one-digit ICSO08 occupation, while Figure 3 capture the distribution of the modal educational achievement (i.e., it is a distribution of an estimated statistic). This explains why Figure 4 tends to give a more nuanced picture, in which the shares of the different levels of education are more dispersed([23]).

([18]) With the implications that: i) all workers with primary or secondary education working in ISCO08 1-3 are undereducated; ii) all workers with primary or secondary education working in ISCO08 4-9 are well matched. In theory, a distinction is also possible between "middle-skilled" occupations (ISCO08 4-8) for which secondary education is appropriate level and "low-skilled" occupations (ISCO08 9) for which the appropriate level of education is primary. In this case also workers with secondary education could be considered overeducated if they work in ISCO08 9 occupations.

([19]) This is done by aggregating at the one-digit ISCO08 level the results obtained for two-digit ISCO08 occupations, using the realized matches method.

([20]) In ISCO08 1 the vast majority (about 83%) of two-digit occupations indeed have tertiary education as their mode. However, for about 16% of occupations the mode is secondary education (a small percentage also have short-cycle tertiary education as their mode). As we move to ISCO08 2, we can see that the realized matches and the Eurostat methods almost fully coincide. However, for ISCO08 3 a major difference appears again: tertiary education (combination of ISCED11 5 and ISCED11 6-9) is the prevailing mode for only about 55% of two-digit occupations in ISCO08 2 (it is 100% according to Eurostat), with a large share of two-digit occupations in which the mode is upper secondary education (in a very small number of cases the mode is post-secondary). In ISCO08 4 (in which, according to Eurostat, all workers should have tertiary education), we still find about 10% of two-digit occupations in which the mode is tertiary education (ISCED11 5-8).

([21]) In ISCO08 5-9 occupations we can see that ISCED 6-8 is never the modal level of education among two-digit occupations; however there are still some very limited cases in which the mode is short-cycle tertiary education (ISCED11 5), in contrast with what is assumed by the Eurostat method.

([22]) As mentioned before, the realized matches approach suffers from endogeneity, in the sense that occupations in which workers tend to have a level of education higher than the one required by the job will tend to have a high level of modal educational attainment. But then, in these occupations, we will end up underestimating the level of overeducation, exactly because there are many workers who are *de-facto* over-educated. This problem, which might affect Figure 3, does not exist for Figure 4.

([23]) In essence, for each two-digit ISCO08 occupation, Figure 3 does not consider all those individuals that have educational attainment levels that are different from the modal one.

Figure 3. Modal level of Education for 1 digit ISCO occupation.
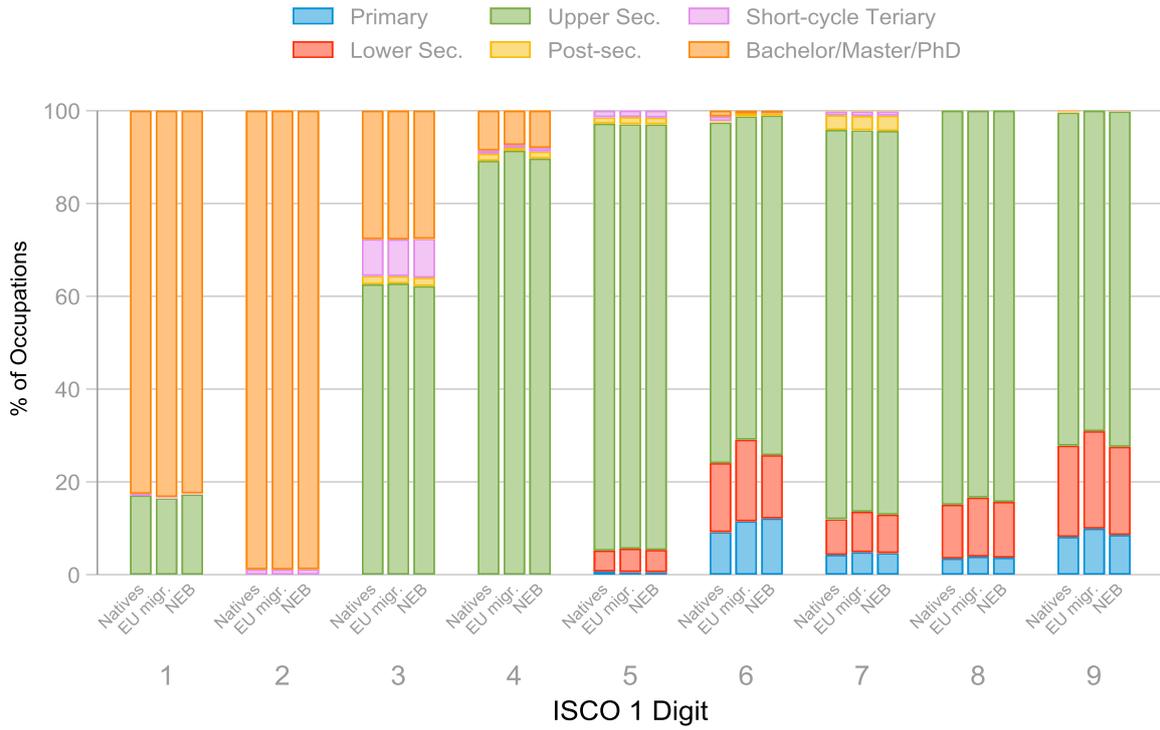
Figure 4. Educational level of workers in ISCO 1 digit occupations

Figure 4 reveals that in one-digit ISCO08 occupations 1, 2 and 3 there are significant shares of individuals with (lower and upper) secondary education (more so among natives), and even some workers with only primary education (more so among NEB) while in one-digit ISCO08 6 to 9 we also observe some individuals with tertiary education (more so for EU migrants). The differences between Figure 3 and 4 are especially evident for ISC08 occupations 1, 2, 4, 5, 6, 7, 8, 9. In any case, this evidence points to the importance of considering within one-digit ISCO variation as well.

3.2 Overeducation among secondary and tertiary graduates

In the remainder of this report we focus on workers with secondary and tertiary educational achievements and drop 196,391 individuals whose highest educational achievement is at most primary education. Therefore, we are left with a working sample of 5,757,908 individuals.

In Figure 5 to Figure 10 we present the percentage of natives, EU migrants and NEB workers who are overeducated according to the realized matches (Figure 5, Figure 6,

Figure **7** and Figure 9) and Eurostat (

Figure **8** Figure 10) definitions for each country. We have added a 45 degree dashed line that allows comparing visually the overeducation level for the two sub-populations of natives and migrants; a dot above the line indicates that natives are more overeducated than migrants while the opposite occurs if the dot is below the line. These are hence providing two sources of information: on the absolute as well as the relative levels of overeducation of the two groups.

3.2.1 Secondary graduates

Here we look at secondary graduates, that is to say, at individuals whose educational achievement is at least lower-secondary (ISCED 2) and at most post-secondary education (ISCED 4), and compare natives with NEBs and EU migrants. We only use the realized matches method.

From Figure 5 we can notice that most EU MS show low values of overeducation (less than 20%) and similar patterns for natives and NEBs. However, southern countries (with Luxembourg and Lithuania) show mid-to-high levels of overeducation among NEB secondary graduates, with large differences relative to natives in Spain and Luxembourg.

A very similar picture emerges when we compare natives to EU migrants (but in this case Lithuania is no longer among the outliers: see Figure 6)

Figure 5. Overeducated by Member State (%) - Natives vs. NEB
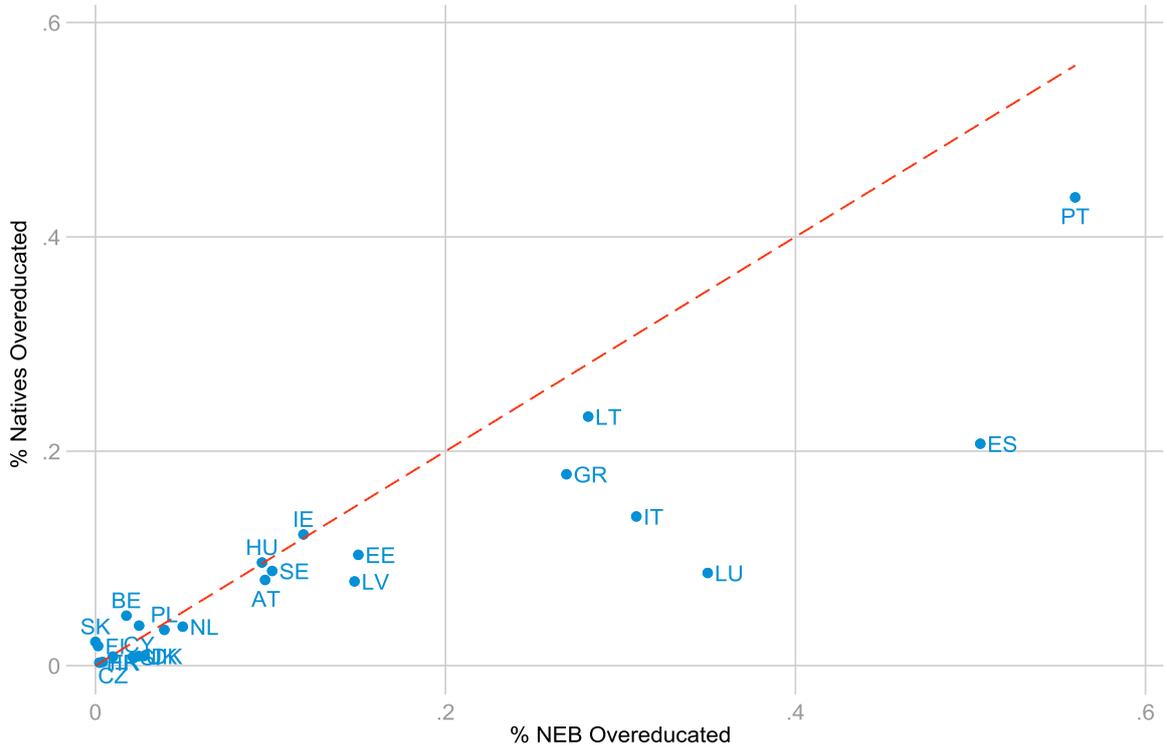
Figure 6. Overeducated by Member State (%) - Natives vs. EU Migrants.
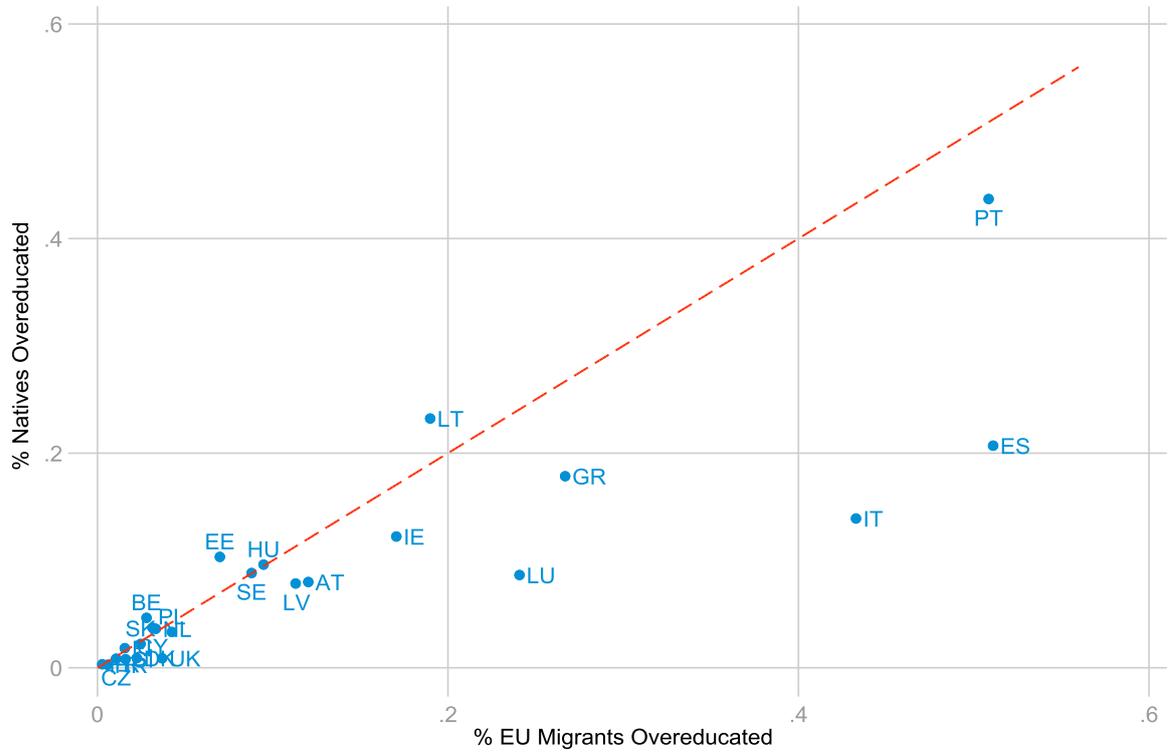
3.2.2 Tertiary graduates

When comparing natives to NEB (and using the realized matches method), from Figure 7 it is evident that, with the only exception of Slovakia, NEBs tend to be relatively more overeducated. The gap is particularly large in Italy and Greece. These two countries do not only show the widest gap, but are also those where the levels of realized overeducation of NEBs are highest. The case of Italy is particularly startling as it has the highest rate of overeducation in the Union among NEBs (80%) and one of the highest among natives (40%). The figure also shows a clear positive correlation between overeducation for the two groups across the EU: in labour markets where natives are not well matched, also NEBs tend to be not well matched (and vice versa).

Figure 8 provides similar information, but uses the definition of overeducation provided by Eurostat. The major difference from Figure 7 is that the number of natives associated to overeducation decrease substantially. Italy and Greece, however, remain outliers (with the addition of Cyprus and Spain).

These Figures confirm that- for tertiary graduates- important differences exists between the realized matches and the Eurostat method.

When we compare natives to EU migrants (using the realized matches approach), Figure 9 reveals a more balanced picture. In 9 MS natives are more likely to be overeducated than EU migrants while the opposite is true in 17 MS (however Hungary and Croatia are very close to the 45-degree line) with Italy, Austria and the UK being characterized by very high levels of overeducation among EU migrants.

The levels of overeducation among EU migrants and the difference with natives tend to be higher in Mediterranean countries (with UK and Ireland) when we use the Eurostat definition (Figure 10).

Figure 7. Overeducated by Member State (%) - Natives vs. NEB (realized matches)
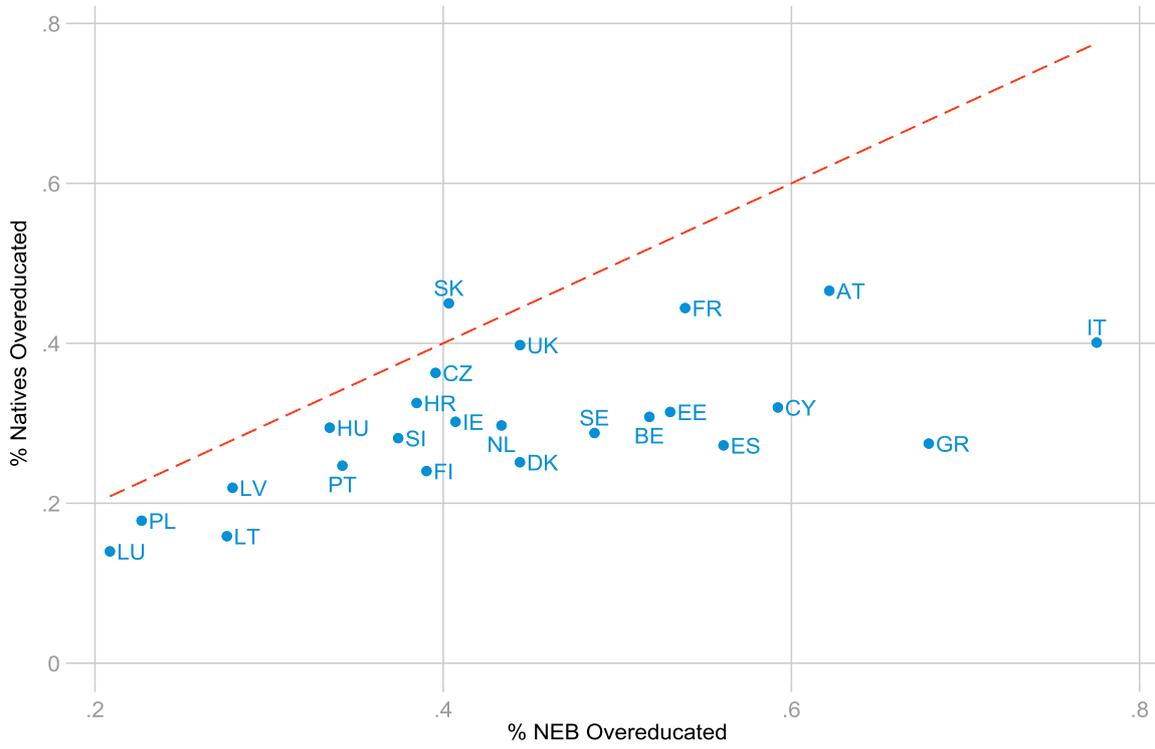
Figure 8. Overeducated by Member State (%) - Natives vs. NEB (Eurostat)
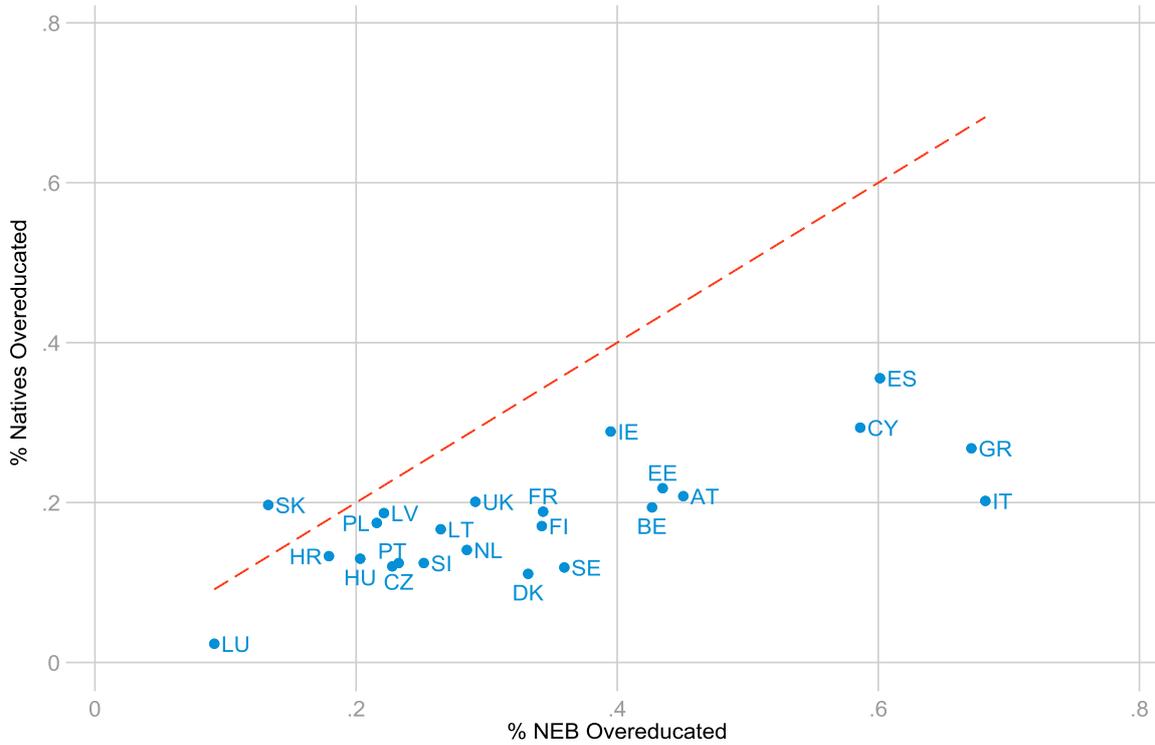
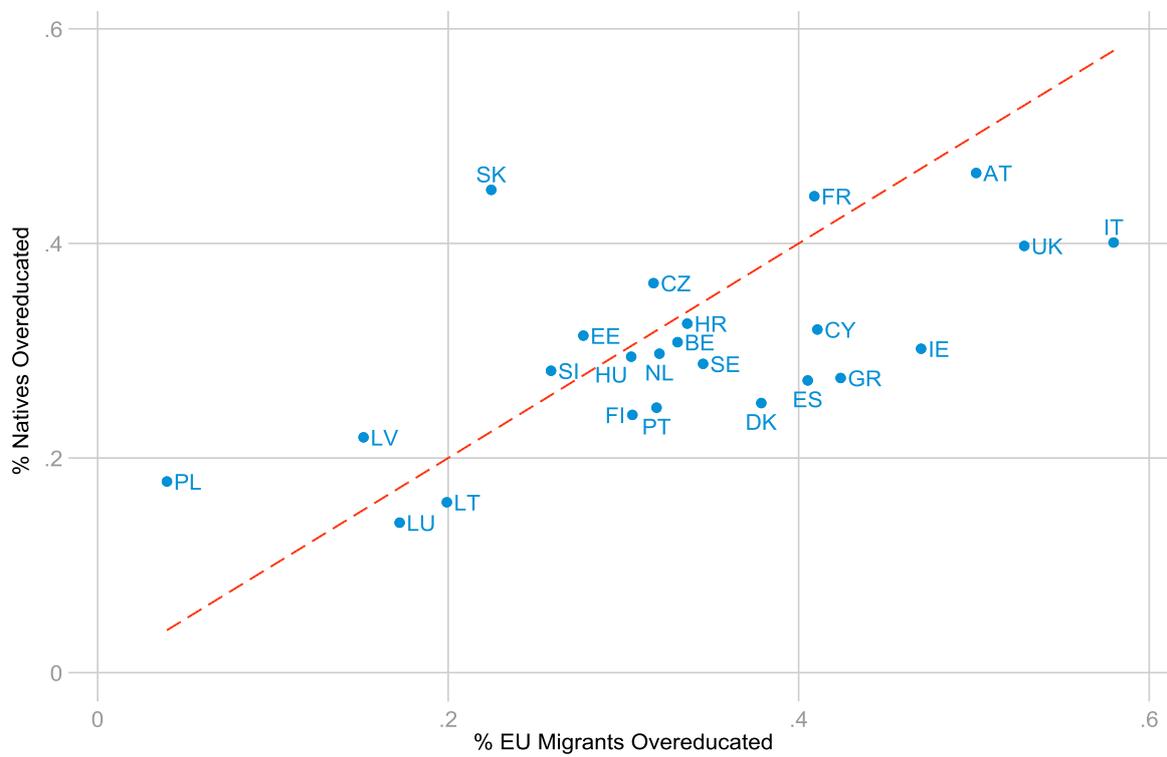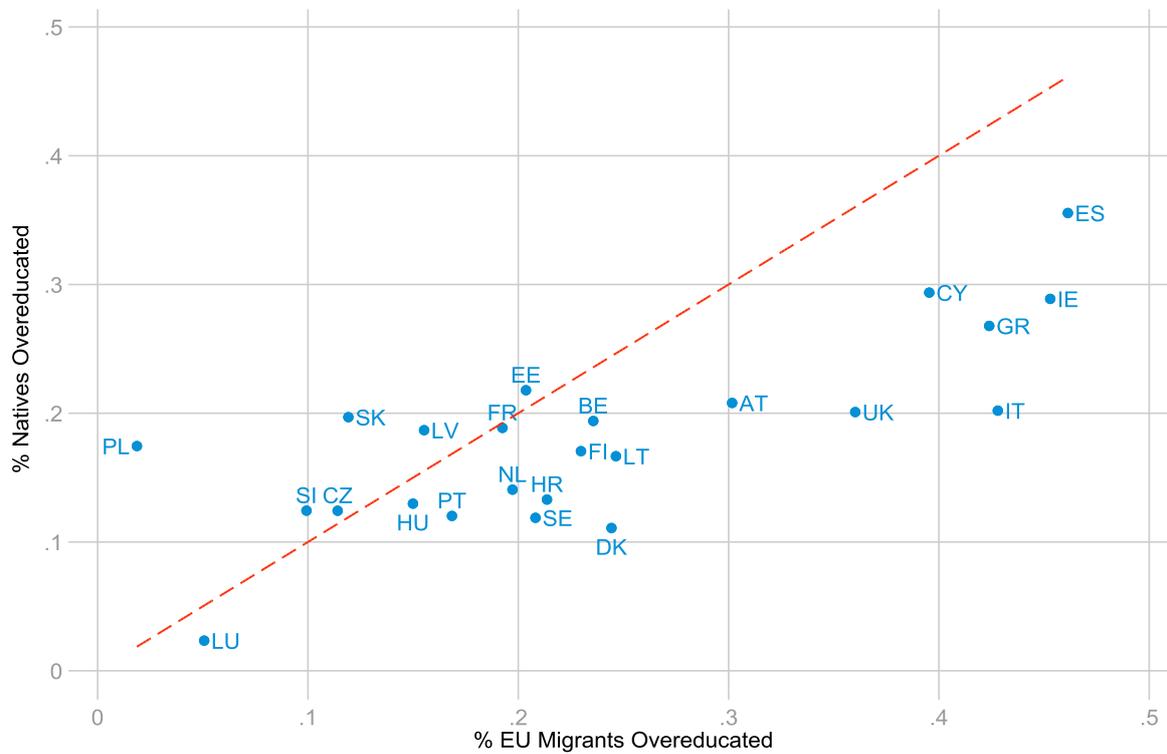**Figure 9.** Overeducated by Member State (%) - Natives vs. EU Migrants (realized matches)

Figure 10. Overeducated by Member State (%) - Natives vs. EU Migrants (Eurostat)

4 Regression analysis

In the previous sections we have looked at over/undereducation from different angles, but also limiting ourselves to basic summary statistics. In this Section we move forward and try to investigate more thoroughly whether the observed patterns change when we impose the *ceteris paribus* condition. This is an especially important point when trying to derive policy implications, since compositional effects might play a decisive role in accounting for the observed patters. For instance, the differences in the raw overeducation rates between EU natives and NEBs that we report in Figure 7-8 might be due to several factors and they do not necessarily reflect either an inefficiency of labour markets in allocating migrants to occupations commensurate to their skill levels, or intrinsic differences in productivity between local and foreign college graduates. In fact, overeducation could be influenced by age. It is reasonable to think that younger individuals entering the labour market might be in search of the right job and until the right match occurs they might accept a job offer that does not fully reward their educational investment. Since, on average, migrants tend to be younger than natives, the higher incidence of overeducation among NEBs could be explained by the different demographic structure of the two groups.

To investigate the gap further, we resort to regression analysis in which observable demographic and other differences between the two groups are accounted for (*ceteris paribus* assumption). As in Section 3, we distinguish between tertiary and secondary graduates and for tertiary graduates we analyse overeducation applying both the Eurostat definition and the realized matches method.

4.1.1 Secondary graduates

We start to present our results by first looking at secondary educated individuals. In Figure 2 we have shown that, within this group, overeducation is particularly widespread among individuals with post-secondary education and fairly uncommon among lower and upper-secondary graduates; at same time, as shown in Figure 1, the vast majority of individuals with completed secondary education are from one of these two categories.

In **Error! Reference source not found.**, we regress workers' match in the labour market on year, country and industry fixed effects, age and its squared term, the degree of urbanization in the area of workers' residence and a dummy for gender. To test whether the matching relationship is affected by the length of stay in the host country, in one specification of this baseline model we distinguish migrants according to their length of stay in the host country (short vs. long-term residents([24])). The intuition here is that migrants might need some time to find the right match in the labour market. This might be due, for example, to migrants' learning on how the labour market functions in the host country (i.e., different mechanisms for job search) and/or to a process of skill acquisition (i.e., language proficiency) that are specific to the labour market of destination.

Our estimates are obtained using an ordered probit where the outcome variable (match status) takes on three values: undereducated, matched and overeducated. In **Error! Reference source not found.** we only present the average marginal effect of the ordered probit, while the coefficients are presented in the appendix.

Column 1 presents the differences in probabilities between natives and migrants of being in one of the three matching statuses for the full sub-sample of secondary educated individuals. On average, EU migrants are 3.8% and 2% more likely to be overeducated and well matched respectively and 5.8% less likely to be undereducated. The picture for NEBs is fairly similar: they are 5% and 2.4% more likely to be overeducated or well

---

([24]) The EU LFS collects information on when the migrant has moved to the country of residence, this allows us to distinguish between long and short term migrants, where we define the former as someone who has been living in the hosting country for less than 5 years and the latter as someone who has been living there for longer than that.

matched and 7.4% less likely to be undereducated than natives. All marginal effects are very precisely estimated and highly statistically significant. In Column 2, we introduce a distinction between short and long-term migrants, but we see that the results confirm those of Column 1 (with the additional indication that overeducation is slightly higher for long-term migrants and for NEB, relative to EU migrants).

Since the post-secondary category is the one presenting the highest incidence of overeducated individuals (see Figure 2), we are concerned that our results might be driven mostly by this category: in Columns 3 and 4 we repeat the same analysis presented in the first two columns, but excluding post-secondary graduates. Clearly, our coefficients are only marginally affected by this sample restriction and we can safely conclude that our results are valid for the whole secondary graduates group.

Table 1. Probability of over-education. Realized matches - Secondary educated individuals

	All		*Excluding Post-secondary*	
	(1)	(2)	(3)	(4)
EU migrants:				
Overeducated	0.038***		0.035***	
	(0.000)		(0.000)	
Matched	0.020***		0.025***	
	(0.000)		(0.000)	
Undereducated	-0.058***		-0.060***	
	(0.000)		(0.000)	
NEB:				
Overeducated	0.052***		0.044***	
	(0.000)		(0.000)	
Matched	0.025***		0.031***	
	(0.000)		(0.000)	
Undereducated	-0.074***		-0.076***	
	(0.000)		(0.000)	
EU recent migrants:				
Overeducated		0.028***		0.027***
		(0.000)		(0.000)
Matched		0.019***		0.025***
		(0.000)		(0.000)

Undereducated		-0.047***		-0.052***
		(0.000)		(0.000)
EU long-term migrants:				
Overeducated		0.040***		0.037***
		(0.000)		(0.000)
Matched		0.022***		0.028***
		(0.000)		(0.000)
Undereducated		-0.062***		-0.063***
		(0.000)		(0.000)
NEB recent migrants:				
Overeducated		0.038***		0.035***
		(0.000)		(0.000)
Matched		0.021***		0.028***
		(0.000)		(0.000)
Undereducated		-0.060***		-0.063***
		(0.000)		(0.000)
NEB long-term migrants:				
Overeducated		0.051***		0.043***
		(0.000)		(0.000)
Matched		0.024***		0.032***
		(0.000)		(0.000)
Undereducated		-0.075***		-0.078***
		(0.000)		(0.000)
N	2,814,358	2,814,358	2,607,884	2,607,884

All regressions include year, country and industry fixed effects and controls for gender, age, degree of urbanization and firm size. *P-values* in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

4.1.2 Tertiary graduates

In Table 3 we report the results of two different specifications for two different econometric models. In Columns 1-4 we report the mean marginal effects[25] of our usual ordered probit model in which we use the realized matches measure for overeducation, while in Columns 5 and 6 we report the main coefficients of a linear probability model in which we follow the Eurostat method to define an individual as overeducated. As previously discussed, the realized matches definition allows us to distinguish three different states: undereducated, properly matched and overeducated, indicating, respectively, individuals whose educational level is lower, equal or higher than the mean level of education for one's occupation. This definition differs from the Eurostat definition that allows for only two states: properly educated and overeducated. All models and specifications account for age (and its square) and gender of the respondent, the size and type of industry of the firm in which he/she is employed, the degree of urbanization of the area of residence and the country of residence and the survey year.

Let's now turn to the interpretation of the coefficients reported in Table 3, starting from Columns 1 to 4. In these specifications we adopt the realized matches approach and distinguish between short-cycle tertiary (ISCED11 5) and long-cycle tertiary education (ISCED11 6-8), since we have documented in Figure 2 that the two groups behave quite differently in terms of over/undereducation. In particular, while for short-cycle tertiary it is possible to define three states (overeducated, undereducated and well matched), for long-cycle tertiary graduates only two states are possible (overeducated and well matched). This also implies that for the latter group we estimate a linear probability model, while for the former we estimate an ordered probit[26] where overeducation is the lower outcome and undereducation the highest.

The coefficients in Columns 1-2 can be interpreted as the average difference, across industries, countries and years, in the probability of being overeducated between EU or NEBs migrants and country natives; in Columns 3-4 the interpretation is similar but we have three possible states.

From Column 1 we can see that –among long-cycle tertiary graduates (ISCED11 6-8)- NEBs are around 4% more likely to be overeducated than similar natives. No significant difference with natives are found, instead, for EU migrants. When we distinguish between short-term and long-term migrants (Column 2) we can see that the overeducation gap between natives and migrants tend to increase with time, and more so for NEBs (i.e., overeducation is more widespread among NEB migrants who have spent more time in the host country).

For ISCED11 5 graduates[27], we see (Column 3) that EU migrants are more likely (6%) to be overeducated and less likely to be well matched or undereducated (relative to natives). Similar results apply to NEBs (10% more likely to be overeducated than natives). When we consider the length of stay in the host country (Column 4) we find that: i) all groups of migrants (EU and NEB, short and long term) are more likely to be overeducated (and less likely to be well matched or undereducated) than natives; ii) overeducation is particularly likely among the long-term NEB migrants who are around 11% more likely than natives to be in that state; iii) long-term EU migrants are less likely to be overeducated relative to short term EU migrants (opposite to what is found for ISCED6-8).

Overall, these results also indicate that short-cycle tertiary and long-cycle tertiary graduates behave quite similarly in terms of over/undereducation, but it needs to be noted that among EU migrants, short-cycle tertiary educated individuals perform worse than their tertiary educated counterparts.

---

[25] The coefficients for the ordered probit are presented in the Appendix.
[26] We have also estimated the same model in a multinomial setting. The results are similar between the two specifications.
[27] For this group we do not have a direct comparison with results obtained using the Eurostat method.

If we focus our attention on ISCED11 6-8 graduates, which are the largest group among tertiary graduates, we can see that our results from the realized matches approach are similar to those observed for secondary graduates: NEBs are more likely to be overeducated, and more so if they are long-term residents of the home country[28].

Lastly, we perform the same analysis adopting the Eurostat approach. In this case we use a linear probability model with only two states (overeducated and well matched), and we regress the dependent variable[29] on the variable capturing the area of origin (EU native, EU migrant, NEB migrant), while controlling for all the characteristics mentioned above. The coefficients in Column 5 and 6 can be interpreted as the average difference, across industries, countries and years in the probability of being overeducated between EU or NEBs migrants and natives. Column 5 refers to a regression in which we do not distinguish by length of stay in the host country, while in Column 6 we take that into account (short vs. long term residents). We can immediately see that, once we include covariates such as age, gender, degree of urbanization and firm's size, NEB and EU migrants are, respectively, around 5% and 3% more likely to be overeducated relative to natives with the same observable characteristics. When we consider the length of stay in the labour market (Column 6), we find that for all groups of NEBs and EU migrants overeducation is more prevailing among long-term residents (and more so for NEBs).

While not referred to the same sample[30], the results from Column 5 are consistent with those of Column 1, while those of Column 2 are consistent with those of Column 6.

Table 2. Probability of over-education. Realized matches and Eurostat definition - Tertiary educated individuals

	Realized matches				Eurostat definition ISCED (5-8)	
	Tertiary (ISCED 6-8)		Short Tertiary (ISCED 5)			
	(1)	(2)	(3)	(4)	(5)	(6)
EU migrants:						
Overeducated	0.002		0.061***		0.032*	
	(0.074)		(0.000)		(0.005)	
Matched			-0.004***			
			(0.000)			
Undereducated			-0.057***			
			(0.000)			
NEB:						
Overeducated	0.039***		0.108***		0.051*	
	(0.000)		(0.000)		(0.000)	
Matched			-0.009***			

---

[28] In Section 4.1.1 we showed that, among secondary graduates, EU migrants and NEBs are more likely to be over-educated

[29] Since we have only two states, the dependent variable takes a value of one when an individual is overeducated and zero otherwise (i.e., when it is well matched).

[30] In Columns 1-4 we distinguish between short term tertiary (ISCEC11 5) and tertiary (ISCED11 6-8), so that the results from the regression that uses realized matches are not fully comparable to those of Column 5 and 6, where the definition of tertiary graduates includes graduates from ISCED11 5 to ISCED11 8.

	(1)	(2)	(3)	(4)	(5)	(6)
			(0.000)			
Undereducated			-0.099***			
			(0.000)			
EU recent migrants:						
Overeducated	0.021***		0.080***		0.037***	
	(0.000)		(0.000)		(0.000)	
Matched			-0.007***			
			(0.000)			
Undereducated			-0.073***			
			(0.000)			
EU long-term migrants:						
Overeducated	0.029***		0.057***		0.045***	
	(0.000)		(0.000)		(0.000)	
Matched			-0.005***			
			(0.000)			
Undereducated			-0.053***			
			(0.000)			
NEB recent migrants:						
Overeducated	0.093***		0.062***		0.047***	
	(0.000)		(0.000)		(0.000)	
Matched			-0.005***			
			(0.000)			
Undereducated			-0.057***			
			(0.000)			
NEB long-term migrants:						
Overeducated	0.128***		0. 114***		0.121***	
	(0.000)		(0.000)		(0.000)	
Matched			-0.011***			
			(0.000)			
Undereducated			-0.103***			
			(0.000)			
N	1,394,348	1,394,348	379,985	379,985	1,563,017	1,563,017
R^2	0.190	0.190			0.207	0.207

All regressions include year, country, field of study and industry fixed effects and controls for gender, age, degree of urbanization and firm size. *P-values* in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

4.2 How to Read our Results

Our results describe a situation in which secondary and tertiary educated migrants- and particularly NEBs- underperform natives on the quality of their labour market match. These results are in line with most literature on overeducation among migrants. We have also documented how the incidence of overeducation is higher among long-term rather than short-term migrants, and particularly so for NEBs. At the same time we find that secondary graduate migrants (both EU migrants and NEBs) are also more likely to be well matched (something we do not find for tertiary graduate migrants), while they are less likely to be undereducated (in common with tertiary graduate migrants).

Our results hence point to the recognition that tertiary graduate migrants – and especially NEBs- do not get full use of their educational attainment. This might be due to difficulties in obtaining formal accreditation of their graduate degrees, to the fact that entry into the labour market for tertiary graduates is particularly difficult, due to barriers or institutional complexities([31]) or that similar degrees in different countries provide different skills and qualifications that the labour market is able to recognize, or to labour market characteristics that end up penalizing tertiary educated migrants (such as the presence for strong social networks). What is particularly worrisome is that the overeducation gap between natives and tertiary educated migrants does not decrease (actually increases) with the length of residence in the home country.

For secondary graduates the picture is more nuanced, as they are more likely to be both overeducated and well matched. This might be due again to difficulties in obtaining accreditation of their educational careers. However, for this group, we expect this problem to be less prevalent. What might be happening here is that the labour market, due to the lower entry barriers and lower controls- is better able to select migrant workers endowed with the appropriate skills. At the same time, we confirm for secondary graduates that overeducation does not decrease with the length of stay in the home country.

More generally, we have to be aware that self-selection of migrants is likely to play a major role and it could operate differently for the two skill groups. Unfortunately, our analysis cannot provide a definitive answer on the ultimate drivers of our results and we should exercise some caution in drawing policy implication from them, for at least two reasons: our data suffers from important limitations and our econometric methodology is not robust to self-selection and thus endogeneity.

Regarding data, there are at least two glaring limitations in the information available in the LFS: a) we have no information on the exact county of origin of the individual; b) we do not know where the individual obtained the highest degree. The coefficients that we present here are an average effect across all countries of origin/education, but it is reasonable to think that important differences exist within the group, as the quality of the match depends on the country of origin/education.

With respect to our methodology, we have estimated our parameters with a fairly standard nonlinear limited dependent variable model. This model is appropriate given our research question, but it does not allow for a causal interpretation of our estimates as the applied method can only account for what is observable, and we have to be aware that relevant differences that could have important implications for our results are not observed (or not observable) in the data at hand.

([31]) This is typically the case for professional figures such as medical doctors, architects, and lawyers.

5 Conclusions

In this report we have analysed in details the extent of educational mismatches in the EU focusing on the different experience of migrants (from the EU and from outside the EU) and EU natives, using LFS data for the years 2011-2016. Specifically, we have considered over/undereducation among the groups of tertiary graduates (ISCED11 level 5 to 8) and secondary graduates (ISCED11 level 2 to 4), comparing natives to EU migrants and NEBs.

The first challenge we faced was that of how to measure over/undereducation. In the literature different approaches have been proposed and we (mostly) relied on the realized matches method, where the level of education of each worker is compared to the one prevailing (i.e., the mode) in the two-digit ISCO08 occupation in which she/he is employed. For tertiary graduates we also used a method proposed by Eurostat, by which tertiary graduates who are working in occupations that do not require such education level are considered as overeducated (i.e., those with educational attainment ISCED11 5 to 8 and employed in occupations in one-digit ISCO08 categories 4 to 9).

When looking at basic statistics on overeducation, our analysis has shown that, for secondary graduates, in the majority of MS the differences between natives, NEBs and EU migrants are not very large, with the exception of few countries –most of them in the south of the EU- in which overeducation is clearly more frequent among migrants.

As for tertiary graduates, larger differences emerge between natives and migrants: overeducation is definitely more frequent among migrants in the vast majority of MS, and particularly so in some southern countries, in which overeducation is also quite high among natives.

Taken at face value, these results would indicate that: i) overeducation is an important phenomenon (especially for some MS located in the south of Europe and more so when we use the Eurostat method); ii) it tends to affect migrants more than natives; iii) it affects tertiary graduates (irrespective of the method used to measure it) more than secondary graduates.

However, these basic descriptive statistics are likely to be affected by how individual characteristics (potentially affecting overeducation) are distributed among native and migrant workers in the sample used in the analysis. To further investigate whether the status of migrant (vs. native) *per se* can affect the likelihood of a worker being overeducated, we have complemented the descriptive statistics with a regression analysis, where we directly control for the observable characteristics that might correlate with the dependent variable. The results obtained show a similar picture. After controlling for year, country and industry fixed effects, age and its squared term, the degree of urbanization in the area of workers' residence and gender, we confirm that EU migrants and – especially-NEBs with secondary and tertiary education are more likely to be overeducated (and less likely to be undereducated) compared to natives with the same educational level. Moreover, our results indicate that for migrants the likelihood of being overeducated increases with the length of stay in the home country, and particularly so for NEBs. This clearly points to a major obstacle to full integration of migrants into EU labour markets. There could be various reasons for this: from barriers to entry (likely to be more stringent for tertiary graduates), to a lower level of usable human capital embedded in migrants, to labour market discrimination. With the data at hand we cannot distinguish between these hypotheses.

Another important result of our analysis has been to show how (and why) the realized matches approach differs from the one proposed by Eurostat for tertiary graduates. In the latter, all workers that have tertiary education are considered as overeducated if they work in (one-digit) ISCO08 4-9 occupations, while they are considered well matched if they work in (one-digit) ISCO08 1-3 occupations. However, we have shown that there are many two-digit ISCO occupations falling in ISCO08 1-3 in which the modal educational attainment is lower than tertiary, and, symmetrically, that there are some two-digit ISCO occupation belonging to ISCO 4-9 in which tertiary education is the mode. Our point here is not to

conclude that one method is better than the other one. In fact, they both have plus and minuses, of which policy makers should be well aware. In principle, the job analyst method (of which the one proposed by Eurostat is a rough approximation) is preferable, as long as it reflects differences across countries and time. We think that the development of ESCO classification is a very important step in this direction and future research should try to use this important source of information for the estimation of vertical and horizontal skill mismatches.

One of the (indirect) goals of this report is to show that both theory and data matter: the way in which we define a phenomenon and in which we collect data to measure it have profound effects on the "big numbers" we obtain at the end of the process (e.g., across MS comparisons of overeducation rates for migrants and natives). In particular, we think that more attention should be given to the problem of selection bias, i.e., to the forces that led us to observe the sample that we use for our analysis. In our case, the problem is particularly important since migration decisions and permanence in the labour market clearly depends upon individuals' characteristics that we cannot control for, making it difficult to draw policy implications –which are by construction based on a cause-to-effect logic- in situations in which we observe only part of the causes and we do not know how observables and unobservable characteristics are correlated

Future research should also address the following issues: 1) focus on "robust" objective measures of vertical (and horizontal) mismatches (ESCO is a source to be explored); 2) use data and methods that are less likely to be affected by selection bias, for instance applying robust econometric methods to longitudinal data; 3) explore the gender dimension of overeducation and how it differs between natives and migrants (with possible implications for family compositions and family reunion); 4) explore the importance of country of origin/destination and country of educational attainment in the likelihood of over/undereducation of migrants; 5) analyse the consequences of over/undereducation (and more generally of skill mismatches) on wages, job satisfaction, job mobility, length of permanence in the host country, and whether overeducation at the beginning of the working experience has a scarring effect, especially for migrants; 6) evaluate the impact of specific policy initiatives –including policies for the selection of migrants- on the labour market integration of migrants using counterfactual methods, so as to derive clear and precise policy implications on what works (and for whom).

References

CEDEFOP, *The skill matching challenge: Analysing skill mismatch and policy implications*, Publications Office of the European Union, Luxembourg, 2010.

Desjardins, R., and Rubenson, K., *An analysis of skill mismatch using direct measures of skills*. OECD Education Working Papers, Paris, 2011.

Dolton P., and Vignoles A., `The incidence and effects of overeducation in the U.K. graduate labour market', *Economics of Education Review*, Vol. 19, No 2, 2000, pp. 179–198.

European Commission, Measuring Skills Mismatch, Analytical Web Note 7/2015.

Eurostat, Looking for immigrants in the European Labour Force Survey and the EU census: a comparison based on the 2011 figures, 2018 Edition, Statistical Report, 2018.

Flisi S., Goglio V., Meroni E.C., Rodrigues M., and Vera-Toscano E., 'Measuring Occupational Mismatch: Overeducation and Overskill in Europe—Evidence from PIAAC', *Social Indicators Research*, Vol. 131, 2017, pp. 1211–1249.

Groot, W., and van den Brink, H. M., 'Overeducation in the labor market: A meta-analysis', *Economics of Education Review*, Vol. 19, No 2, 2000, pp. 149–158. doi:10.1016/S0272-7757(99)00057-6.

Hartog, J., 'Over-education and earnings: Where are we, where should we go?', *Economics of Education Review*, Vol. 19, 2000, 131–147.

Hartog, J., and Oosterbeek, H., 'Education, allocation and earnings in the Netherlands: Overschooling?', *Economics of Education Review*, Vol. 7, No 2, 1988, pp. 185–194.

Kiker, B. F., Santos, M. C., and Mendes De Oliveira, M., 'Overeducation and undereducation: Evidence for Portugal', *Economics of Education Review*, Vol. 16, No 2, 1997, pp. 111–125.

Leuven, E., and Oosterbeek, H., 'Overeducation and mismatch in the labor market', In E. Hanushek, S. Machin and L. Woessmann, eds. *Handbook of the Economics of Education*, Elsevier Science, Vol.. 4, 2011, pp. 283-326.

McGoldrick, K., and Robst, J., 'Gender differences in overeducation, a test of the theory of differential overqualification'. *American Economic Review*, Vol. 86, No 2, 1996, pp. 280-284.

McGuinness, S., 'Overeducation in the labour market', *Journal of Economic Surveys*, Vol. 20, No. 3, 2006, pp. 387–418. doi:10.1111/j.0950-0804.2006.00284.x.

McGuinness, S., Pouliakas K., and Redmond P., Skill mismatch: concepts, measurement and policy approaches, *Journal of Economic Surveys*, Vol. 32 No.4, 2018, pp. 985-1015. doi: 10.1111/joes.12254

Meroni E.C. and Vera-Toscano E., 'The persistence of overeducation among recent graduates', *Labour Economics*, Vol. 48, 2017, pp. 120-143.

OECD, *Better Skills, Better Jobs: A Strategic Approach to Skills Policies*, OECD Publishing, Paris, 2012. http://dx.doi.org/10.1787/9789264177338-en.

OECD/European Union, *Matching Economic Migration with Labour Market Needs*, OECD Publishing, Paris, 2014. *http://dx.doi.org/10.1787/9789264216501-en*

OECD/ European Union, *Settling In 2018: Indicators of Immigrant Integration*, OECD Publishing, Paris/EU, Brussels, 2018, *https://doi.org/10.1787/9789264307216-en*.

Pellizzari, M., Biagi F., and Brecko B*., E-skills Mismatch: Evidence from International Assessment of Adult Competencies (PIAAC)*, JRC Technical Reports, Institute for Prospective Technological Studies, Digital Economy Working Paper 2015/10, 2015.

Pellizzari, M., and Fichen, A., *A new measure of skills mismatch*, OECD Social, Employment and Migration Working Papers, Paris, 2013. http://www.oecd-ilibrary.org/content/workingpaper/5k3tpt04lcnt-en.

Quintini, G., *Over-qualified or under-skilled*. OECD Social, Employment and Migration Working Papers. Paris, 2011. http://www.oecd-ilibrary.org/content/workingpaper/5kg58j9d7b6d-en.

Rumberger, R., 'The impact of surplus schooling on productivity and earnings', *Journal of Human Resources*, Vol.. 22, No 1, 1987, pp. 24-50.

Verdugo, R. R., and Verdugo, N. T., 'The impact of surplus schooling on earnings: Some additional findings', *Journal of Human Resources*, Vol. 24, No 4, 1989, pp. 629. doi:10.2307/145998.

Verhaest, D., and Omey, E., 'Measuring the incidence of over- and undereducation'. *Quality and Quantity*, Vol. 40, 2006, pp. 783-803.

Verhaest, D., Sellami S., and van der Velden, R., Differences in horizontal and vertical mismatches across countries and fields of study, International Labour Review, Vol. 156 (2017), No.1, pp. 1-23.

Visintin, S., Tijdens, K., and van Klaveren, M., 'Skill mismatch among migrant workers: evidence from a large multi-country dataset'. *IZA Journal of Migration*, Vol. 4, No 14. DOI: 10.1186/s40176-015-0040-0

List of abbreviations and definitions

EU   European Union

ESCO   European Skills, Competences, Qualifications and Occupations classification

ISCED   International Standard Classification of Education

ISCO   International Standard Classification of Occupations

MS   Member State

OECD   Organization for Economic Co-operation and Development

NEB   Non-EU born

List of figures

List of tables

Annexes

Annex 1. Job Match by Region of Origin

If we compute the shares of overeducated, well matched and undereducated workers, distinguishing between <u>natives</u> and <u>non-EU born</u> (A 1, columns 1-2-3), we immediately notice that in most (i.e., 21) countries NEBs tend to be overeducated when compared to natives (exceptions are Croatia and Slovenia, where the opposite is true). On the other hand, natives are more likely than NEBs to be well matched in 22 MS (the exception is Ireland, where the opposite is true). When considering undereducation, the pattern is much less clear, as there are 15 countries in which non-EU born are more likely to be undereducated relative to nationals, while in other 6 the opposite is true (in the remaining MS the differences are not significant).

When comparing <u>natives</u> to <u>EU migrants</u> (A 1, columns 1-4-5) the picture that emerges is slightly less clear. Overeducation among EU migrants is more prevalent than among natives in 12 MS, but the reverse is true in 4 MS (France, Poland, Slovenia and Slovakia). Natives are more likely to be well matched in 17 MS (the only MS where the reverse is true is Poland), while undereducation is more frequent among natives in 9 MS and among EU migrants in 7 MS (in the remaining MS the differences are not significant).

A 1. Shares for over-educated, well matched and under-educated: Natives vs NEB vs EU migrants (secondary and tertiary education)

	Natives	NEB	Diff.	EU	Diff.
	(1)	(2)	(1)-(2)	(3)	(1)-(3)
Austria					
Overeducated	0.182	0.190	-0.007***	0.246	-0.064***
Matched	0.637	0.523	0.113***	0.620	0.016***
Undereducated	0.181	0.287	-0.106***	0.134	0.047***
Belgium					
Overeducated	0.170	0.219	-0.050***	0.178	-0.008*
Matched	0.604	0.535	0.070***	0.612	-0.008
Undereducated	0.226	0.246	-0.020***	0.210	0.016***
Cyprus					
Overeducated	0.187	0.240	-0.053***	0.189	-0.002
Matched	0.650	0.494	0.155***	0.626	0.024***
Undereducated	0.163	0.266	-0.103***	0.185	-0.022***
Czech Republic					
Overeducated	0.0840	0.119	-0.035**	0.0952	-0.011
Matched	0.831	0.731	0.101***	0.792	0.040***

Undereducated	0.0847	0.151	-0.066***	0.113	-0.028***
Denmark					
Overeducated	0.123	0.222	-0.100***	0.233	-0.111***
Matched	0.721	0.621	0.099***	0.670	0.050***
Undereducated	0.157	0.156	0.000	0.0963	0.060***
Estonia					
Overeducated	0.182	0.306	-0.124***	0.187	-0.005
Matched	0.591	0.554	0.036***	0.634	-0.043
Undereducated	0.228	0.139	0.088***	0.179	0.049
Spain					
Overeducated	0.246	0.497	-0.252***	0.452	-0.206***
Matched	0.587	0.417	0.171***	0.436	0.151***
Undereducated	0.167	0.0859	0.081***	0.112	0.055***
Finland					
Overeducated	0.132	0.161	-0.029**	0.118	0.014
Matched	0.665	0.590	0.075***	0.560	0.105***
Undereducated	0.204	0.249	-0.045***	0.323	-0.119***
France					
Overeducated	0.176	0.212	-0.036***	0.160	0.016***
Matched	0.604	0.515	0.089***	0.556	0.048***
Undereducated	0.220	0.273	-0.053***	0.284	-0.065***
Greece					
Overeducated	0.216	0.346	-0.129***	0.315	-0.098***
Matched	0.591	0.395	0.196***	0.494	0.097***
Undereducated	0.192	0.259	-0.067***	0.191	0.001
Croatia					
Overeducated	0.0913	0.0792	0.012*	0.110	-0.018
Matched	0.792	0.741	0.051***	0.774	0.019

Undereducated	0.116	0.179	-0.063***	0.116	-0.000
Hungary					
Overeducated	0.140	0.172	-0.031***	0.148	-0.007
Matched	0.723	0.692	0.0315**	0.730	-0.006
Undereducated	0.136	0.137	-0.000	0.123	0.014**
Ireland					
Overeducated	0.222	0.326	-0.104***	0.329	-0.107***
Matched	0.493	0.505	-0.012**	0.471	0.022***
Undereducated	0.285	0.169	0.116***	0.201	0.084***
Italy					
Overeducated	0.192	0.369	-0.177***	0.435	-0.244***
Matched	0.636	0.517	0.119***	0.458	0.178***
Undereducated	0.173	0.114	0.058***	0.107	0.065***
Lithuania					
Overeducated	0.212	0.292	-0.080***	0.208	0.004
Matched	0.629	0.577	0.051***	0.627	0.002
Undereducated	0.159	0.131	0.029***	0.165	-0.006
Luxemburg					
Overeducated	0.106	0.263	-0.158***	0.206	-0.100***
Matched	0.697	0.609	0.088***	0.645	0.052***
Undereducated	0.197	0.128	0.069***	0.149	0.048***
Latvia					
Overeducated	0.129	0.188	-0.060***	0.121	0.007
Matched	0.687	0.691	-0.003	0.653	0.034
Undereducated	0.184	0.121	0.063***	0.226	-0.042*
Netherlands					
Overeducated	0.146	0.178	-0.033***	0.162	-0.017*
Matched	0.643	0.607	0.036***	0.632	0.011

Undereducated	0.211	0.215	-0.003	0.206	0.005
Poland					
Overeducated	0.0861	0.133	-0.047***	0.0438	0.042**
Matched	0.787	0.767	0.020	0.838	-0.051*
Undereducated	0.126	0.0997	0.027*	0.118	0.008
Portugal					
Overeducated	0.359	0.467	-0.108***	0.448	-0.089***
Matched	0.505	0.434	0.071***	0.458	0.047***
Undereducated	0.136	0.0983	0.038***	0.0941	0.042***
Sweden					
Overeducated	0.174	0.275	-0.101***	0.217	-0.043***
Matched	0.643	0.535	0.108***	0.605	0.038***
Undereducated	0.183	0.190	-0.007***	0.178	0.005
Slovenia					
Overeducated	0.108	0.0545	0.054***	0.0803	0.028***
Matched	0.715	0.672	0.043***	0.720	-0.005
Undereducated	0.177	0.274	-0.097***	0.200	-0.023**
Slovak Republic					
Overeducated	0.119	0.169	-0.050**	0.0735	0.045***
Matched	0.814	0.788	0.026	0.828	-0.014
Undereducated	0.0670	0.0430	0.024	0.0980	-0.031**
UK					
Overeducated	0.178	0.277	-0.099***	0.280	-0.102***
Matched	0.592	0.441	0.152***	0.481	0.111***
Undereducated	0.230	0.282	-0.052***	0.239	-0.009

Note: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Notice that in A 1 we have combined workers with secondary and tertiary educational attainment, and the results reflect various compositional effects, both in terms of educational attainment and in labour market participation by EU citizens (natives and migrants) and NEB.

Annex 2. Ordered Probit Estimates

A 2.Ordered Probit Estimated Coefficients.

	Secondary Educated				Short Tertiary	
	All	No Post-sec.	All	No Post-sec		
	(1)	(2)	(3)	(4)	(5)	(6)
Origin:						
EU migrant	-0.447***	-0.531***			-0.014	
	(0.000)	(0.000)			(0.907)	
NEB	-0.318***	-0.376***			0.685***	
	(0.000)	(0.000)			(0.000)	
Female	0.002	0.000			-0.141***	
	(0.726)	(0.997)			(0.000)	
EU migrant X Female	0.085***	0.116***			-0.020	
	(0.000)	(0.000)			(0.667)	
NEB X Female	0.047**	0.068***			-0.046	
	(0.008)	(0.000)			(0.219)	
Age	0.012***	0.007*	0.013***	0.008*	0.046***	0.051***
	(0.000)	(0.019)	(0.000)	(0.011)	(0.000)	(0.000)
Age sq.	-0.000	0.000	-0.000	-0.000	-0.000***	-0.000***
	(0.134)	(0.897)	(0.086)	(0.935)	(0.000)	(0.000)
EU migrant X Age	0.003*	0.004**			-0.002	
	(0.045)	(0.010)			(0.449)	
NEB X Age	-0.002*	-0.002			-0.018***	
	(0.050)	(0.161)			(0.000)	
Firm size	0.034***	0.034***	0.034***	0.034***	0.074***	0.073***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Upper-secondary educ.				-2.223***		
				(0.000)		
Post-sec. non tert. educ.	-3.516***		-3.516***			
	(0.000)		(0.000)			
Origin and length of stay:						

EU recent			-0.165	-0.236		0.140
			(0.191)	(0.074)		(0.528)
EU long-term			-0.555***	-0.638***		-0.337*
			(0.000)	(0.000)		(0.029)
NEB recent			-0.103	-0.169		1.106***
			(0.435)	(0.217)		(0.000)
NEB long-term			-0.362***	-0.418***		0.280*
			(0.000)	(0.000)		(0.018)
cut1	-2.410***	-2.734***	-2.396***	-2.720***	3.083***	3.179***
cut2	0.415***	0.318***	0.428***	0.331***	3.347***	3.443***
Observations	3,046,466	2,826,884	3,046,466	2,826,884	294,951	294,951
chi2	178,442.642	174,142.535	178,468.060	174,165.009	8,038.844	8,120.284

All regressions include year country and industry fixed effects. P-values in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub – Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office