Measuring overeducation using skills requirements

Florian Fouquet

Lemna
EA 4272

Document de travail du LEMNA N° 2023-05

Septembre 2023

Measuring overeducation using skills requirements

Florian FOUQUET

Measuring overeducation using skills requirements

Florian FOUQUET[1]

Abstract

Overeducation can be measured through several approaches, each one leading to different results as to its incidence. This paper aims to build a new and up-to-date measure of overeducation, the *Skill Analysis* measure, based on the content of jobs in terms of tasks and skills required to perform a job. To this end, I mobilise data from the Occupational Information Network (O*NET) and information about learning outcomes from the European Qualifications Framework. I then compare the new measure of overeducation to usual statistical measures using the French Labour Force Survey. Results confirm that objective and statistical measures do not apprehend the exact same matter, showing only moderate correlations between approaches.

Keywords: overeducation, educational mismatch, skills requirements, measurement issues

JEL classification: I21, I23, J24

[1] LEMNA, Nantes Université ; Mail: florian.fouquet@univ-nantes.fr

1. Introduction

From the labour market perspective, education is not always used to its full potential, as some individuals study longer than they actually need to do their job: they are "overeducated" (Kucel, 2011; McGuinness, 2006). According to the European skill and jobs survey (ESJS), about 30% of the adult population of the European Union is mismatched in some way (Cedefop, 2018).

If the increase of overeducation is a widely shared finding in the literature, its incidence is however still debated, especially since it can be measured through several approaches. Capsada-Munsech (2019) show that, depending on the measure, overeducation rates range from 12% to 38% using the same data. The author point that it has long not been of concern because the differences in measures do not substantially affect the returns of overeducation (McGuinness, 2006). However, recent studies discuss its drivers (Croce & Ghignoni, 2012; Davia et al., 2017; Ghignoni & Verashchagina, 2014), using overeducation as a dependent variable and raising the question of what is exactly measured and how.

In this paper, "overeducation"[2] is addressed from the labour market point of view, as an inadequacy between the formal educational level of the individuals and the one required for their job[3]. Overeducation can be seen as an educational investment that did not pay off. There is an opportunity cost to study, because of the direct cost of education and the lack of wages during the education years (Becker, 1964). If additional years of education do not lead to a higher professional position and higher wages, there is a risk that the opportunity cost will never be compensated. This is consistent with both the Human Capital Theory (Becker, 1964) and job-market signalling (Spence, 1973). In the former, education allows individuals to accumulate skills and knowledge that will be used on the labour market; in the latter, educational credentials indicate the level of ability of individuals, that allows them to access better (paid) jobs. In both cases, education is expected to improve labour market outcomes.

Overeducation can be measured through several approaches, each one leading to different results as to its incidence. Statistical measures are based on the distribution of education in the population. Subjective measures are based on workers' assessment about their job situation. Objective measures of overeducation are based on an *ex ante* definition of what should be the "normal" (in the sense of "conforming to a standard")[4] educational level(s) to hold a particular job. Objective measures are often presented as better measures than their statistical or subjective alternatives (Capsada-Munsech, 2019; Verhaest & Omey, 2006).

[2] Overeducation has to be distinguished from other closely related but different concepts: overskilling and horizontal mismatch. Overskilling occurs when the skills of an individual are not fully used in the job she or he occupies. How these skills were acquired (formal education or professional training) is not taken into account, by contrast with overeducation (Capsada-Munsech, 2017). Allen and van der Velden (2001) show that educational mismatch and skill mismatch, although closely related, do not systematically overlap. Horizontal job-education mismatch occurs when individuals do not work in their field of study.

[3] In this regard, the word "overschooling" might seem less ambiguous and a good alternative to the term "overeducation". However, although it has sometimes been used, it remains quite rare in the literature.

[4] This kind of measure is sometimes referred to as "normative measure" in earlier works.

However, they also raise some issues. First, they require to be updated regularly to remain relevant. Due to the cost of keeping them consistent with the evolutions of the labour market structures, these measures are either missing in several countries, or are based on old occupational dictionaries, which raises concerns regarding the accuracy of the measure (van der Meer, 2006). Furthermore, they may not be the most suitable measure for international comparison because of the differences in education and jobs between countries[5].

The aim of this paper is to propose an objective measure of overeducation, built on the skills requirements in each occupation and the learning outcomes of each educational level in Europe, henceforth *Skill Analysis* measure. To this end, I mobilise data from the Occupational Information Network (O*NET) of the US Department of Labor, which provide detailed information about the skills used in each occupation, combined with information about learning outcomes (that is the skills and knowledge provided by each educational level) from the European Qualifications Framework (EQF).

The *Skill Analysis* measure is based on the skills required to do a particular job, with the requirements being defined *ex ante* by experts, and those provided by degrees. More specifically, I first match EQF educational levels to occupations based on the skills scores drawn from O*NET data and then I match these educational requirements to ISCO-08[6] occupations using the correspondence with O*NET. This measure of overeducation relies on the definition of precise occupations to deal with the aggregation problem that arises with the statistical alternatives, *i.e.* the averaging of different educational requirements within larger occupational groups. Moreover, I define two versions of the requirements (upper requirement and lower requirement) to take into account the variety of jobs within a same occupation.

After building the *Skill Analysis* measure, I compare it to existing measures of overeducation, in particular statistical measures. The upper *Skill Analysis* measure leads to the lowest overeducation incidence of all approaches, while the lower *Skill Analysis* shows an incidence similar to what is observed with statistical alternatives. The analysis of the correlations and overlaps between the *Skill Analysis* and the statistical measures shows that the various approaches do not take into account the same individuals, and tend to apprehend slightly different matter.

This paper contributes to the literature about the measure of overeducation, proposing a new measure that deals with the limitations of usual approaches. Other alternative measures have been proposed, using wages (Nauze-Fichet & Tomasini, 2002) or combining several types of measures (Chevalier, 2003). The *Skill Analysis* measure differs from these approaches by computing educational requirements in European countries based on large data about skills and a detailed nomenclature of occupations.

[5] However, in the last decade, there has been an attempt to normalise education systems in Europe, which theoretically reduces differences in educational requirements between those countries.
[6] International Standard Classification of Occupations 2008.

This paper is structured as follows. Section 2 exposes the existing measures of overeducation. Section 3 describes the construction of the *Skill Analysis* measure. Section 4 compares the new measure to existing measures, focusing in particular on young workers. Section 5 concludes.

2. Measuring overeducation

Back in the 1970s and 1980s, the first papers about overeducation relied on an objective measure (Rumberger, 1981). Objective measures are based on a correspondence grid between education and occupations, built *ex ante* by experts according to the content of jobs and degrees. Verhaest and Omey (2006) refer to this measure as *Job Analysis*[7]. Given that the educational norm is defined *a priori*, the objective measure has the advantage not to be affected by the actual distribution of education in the population, which is likely to bias the estimated incidence of overeducation. If this approach has the advantage of being more rigorous, it is however costly. Indeed, it implies the construction and the regular update of a correspondence between education and occupations, in order to keep up with labour market changes and modifications of the content of degrees and jobs.

Statistical measures (also called empirical measures) define "normal" educational levels from the statistical distribution of education in each occupation. These measures can be built from the deviation from the mean (Croce & Ghignoni, 2012; Verdugo & Verdugo, 1989) or from the median, using the mode of the distribution (Davia et al., 2017; Ghignoni & Verashchagina, 2014) or the analysis of contingency tables between jobs and degrees (Forgeot & Gautié, 1997; Nauze-Fichet & Tomasini, 2002). They are easier to use because the required variables to build the measure (the educational level and the occupation of individuals) are available in a large number of datasets. However, as they are based on statistical distributions, they depend on the structure of education in the population studied. In particular, if many workers in an occupation are employed below their educational level, their situation will appear as "normal", although they are objectively overeducated. Verhaest and Omey (2006) refer to these measures as *Realised Matches*. They illustrate the role of education as a signal sent to recruiters. However, as with the objective measures, these approaches are sensitive to the evolution of the structure of jobs and education across time: if educational requirements increase, overeducated workers at a point in time may no longer be mismatched in subsequent periods, without any actual changes in their situation.

Subjective measures are based on the workers' evaluation of their own job situation. Verhaest and Omey (2006) refer to these measures as *Worker Self-Assessments*. Two different

[7] Since it is built by looking at the task content of the jobs, to some extent, this type of measure brings together overeducation and overskilling. However, it does not take into account skills that could have been obtained through professional experience.

approaches can be distinguished, depending on the question asked to respondents. The first approach (direct approach) consists in asking directly to workers if they think they are overeducated. This strategy leads to important biases. First, a similar situation can be interpreted differently by different respondents (one worker declaring being overeducated while another in the same job with the same education states not being overducated). Second, some individuals may choose (deliberately or not) to give a biased answer to understate an overeducation situation (Capsada-Munsech, 2019). Third, overeducation can be mixed up with other job characteristics, such as wage or job satisfaction (Capsada-Munsech, 2019). The second approach (indirect approach) consists in asking workers which degree is required to get or to do their job, referring to signalling (Spence, 1973) or to human capital (Becker, 1964) respectively. While the indirect subjective measure is less biased than the direct approach, individuals may still overstate the importance of their own degree (Capsada-Munsech, 2019). Both approaches rely on unusual data (the answer to the question "Which degree do you need to do/get your job?"), but subjective measures allow to explore how overeducation situations are experienced by workers (Di Paola & Moullet, 2009, 2012), by contrast with objective or statistical measures.

Finally, Chevalier (2003) combines subjective workers' assessment on job satisfaction with an objective definition of educational requirements to build an alternative measure of overeducation, aiming to account for the skill heterogeneity of the workers. The author defines as "genuinely overeducated" individuals who have education in excess of what is required for their job and are dissatisfied at work, and as "apparently overeducated" individuals who have education in excess but are satisfied in their job. Using this methodology, he shows that skills play a role in the risk of overeducation[8].

Because of the variety of approaches to measure overeducation, several studies have explored their reliability and performance. Verhaest and Omey (2006), on Belgian data, compare five measures of overeducation: one *Job Analysis* measure, two *Self-Assessment* measures (depending on the degree needed to do or to get the job) and two *Realised Matches* measures (based on the mean and the mode of the distribution). Using statistical tests (F-tests and LR-tests), they show that the objective measure is better than the others. Indeed, the *ex ante* classification is not impacted by sample characteristics or workers' subjectivity, and should be preferred whenever possible. In a meta-analysis of papers about overeducation from the 1970s to the 1990s, Groot and Maassen van den Brink (2000) find that objective and subjective measures lead to similar overeducation rates. Only the statistical measures showed slightly lower results. More recently, Capsada-Munsech (2019) compare two objective, two statistical and one subjective measures using REFLEX/HEGESCO[9] and LFS data for various European countries, and find low correlations between the different types of approaches. The

[8] Meroni and Vera-Toscano (2017) find that both types of overeducation lead to the same detrimental effects on career paths.

[9] Research into Employment and professional/Higher Education as a Generator of Strategic Competences.

author also underlines the importance of the thresholds above which individuals are overeducated. Finally, for national level analyses, objective measures are the most appropriate, or statistical measures if there is no up-to-date job-education correspondence. For cross-countries comparisons, subjective measures are more suitable because the other approaches raise the question of the comparability of skill and job structures between countries (Capsada-Munsech, 2019).

However, measures do not necessarily substitute for each other. Some individuals are overeducated according to one measure, but are not according to another. This is particularly striking between subjective and objective measures: around 30% of the individuals are overeducated in only one of these two dimensions (Di Paola & Moullet, 2012; Capsada-Munsech, 2019). Thus, using several measures conjointly can bring light on different aspects of overeducation, but the choice of the measure is often driven by data availability (Capsada-Munsech, 2019).

3. Building an objective measure of mismatch based on skills requirements: the *Skill Analysis* measure

3.1. The European Qualifications Framework (EQF)

In order to build an up-to-date objective measure of overeducation, I mobilise information from the European Qualifications Framework (EQF) from the European Commission. The EQF was adopted in 2008 and revised in 2017[10]. It provides information about the qualifications acquired in the different educational systems and labour markets in Europe. The EQF is part of the Bologna Process, which seeks to build a common base for higher education systems in Europe. The aim of this process is to facilitate employability, as well as student and professional mobility.

The EQF defines eight levels of qualifications[11] on a common basis for each member country, with a description of the skills (cognitive use of logical thinking or practical use of methods and tools), knowledge (theoretical learnings) and responsibilities (ability to apply knowledge and skills autonomously) related to each level (European Commission, 2017). Degrees and professional certifications are associated with EQF levels based on the knowledge and skills they provide. EQF levels, the corresponding degrees and the corresponding learning outcomes are described in Table 7 in Appendix A.

[10] See the Council Recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning (European Commission, 2017).
[11] EQF levels 1 and 2 are associated with basic knowledge and skills and correspond respectively to Kindergarten and Primary School in terms of education. They are not levels that can be valued *per se* on the labour market; thus, they are not used in this work.

For the construction of the *Skill Analysis* measure, I gather EQF levels 7 and 8 (Master's degrees and PhD, respectively) for two main reasons. First, based on the descriptions of EQF levels, it is hard to define different scores for levels 7 and 8 in terms of importance and level. Second, in most analyses of overeducation, PhDs are associated with Master's degrees because there are too few PhD graduates to have a reliable sample size. Thus, it should not be of concern not to take EQF level 8 explicitly into account, even if I will not be able to consider PhD graduates working in EQF 7 occupations as overeducated.

3.2. The O*NET database

I also rely on the O*NET (Occupational Information Network) database which is developed by the US Department of Labor and contains detailed information about the content of jobs. It aims at allowing job seekers and recruiters to better understand the labour market, in order to help the former to find the job and the latter to find the workforce that they need. In particular, for each of the 873 occupations in the dataset, the O*NET database provides information about activities and tasks performed, knowledge and skills required, working conditions, experience and education required, and evolution of wages and hirings. To determine whether individuals are overeducated or not, I rely on the information about the use of skills in each occupation.

35 skills are listed in the O*NET data and are grouped into seven categories: Content, Process, Social Skills, Technical Skills, Systems Skills, Resource Management Skills and Complex Problem Solving. Based on the content of occupations, each skill for each of the O*NET occupations is rated by labour analysts on the ground of its importance (1 to 5) and of its required level to perform an occupation (1 to 7)[12]. Ratings are then standardised on a scale ranging from 0 to 100[13]. A score of 100 in terms of importance corresponds to an extremely important skill for the occupation, and a score of 0 to an unimportant skill. A score of 100 in terms of level corresponds to requiring an expert level, and a score of 0 to an unneeded skill. Even though the O*NET database and the EQF come from different countries, the skills required to perform the tasks should not vary substantially for similar jobs, provided that there is no significant technology gap between those countries[14].

[12] The analysts' selection process, as well as a more detailed overview of the notation process, are provided in Fleisher and Tsacoumis (2012).

[13] Standardisation is given by the equation: $S = \frac{O-L}{H-L} \times 100$, with S the standardised score, O the original score given by the analysts, L the lowest possible score on the rating scale and H the highest possible score on the rating scale. For instance, the importance rating scale goes from 1 to 5, so $L = 1$ and $H = 5$. Thus, a standardised score of $S \geq 50$ corresponds to an original score from 3 to 5.

[14] Alternatively, I could rely directly on the information about the degrees required reported in the O*NET database. However, the labour market and education structures are different in the United States than in Europe. Since these institutions strongly depend on political decisions and thus are likely to vary between countries (especially when they do not share a common educational framework), required educational levels in the US may not be the ones required in Europe.

I keep 20 skills out of the 35 skills available. Some skills are excluded because they are not linked to secondary or tertiary education, but to basic knowledge which are not provided by the corresponding degrees. These are the skills of the Content skill category of O*NET[15]. The skill Programming is excluded because it relates to only particular fields (rather than levels) of education and occupations. Finally, I exclude most of the Social Skills and Resource Management Skills[16] because there is no information about them in the EQF. The 20 skills that I use in this work are reported in Table 8 in Appendix B, as well as the description provided by O*NET for each of them.

3.3. Matching skills, degrees and occupations

Using O*NET data and the EQF description of learning outcomes, I aim to define the theoretical education-skills requirements for each ISCO-08 occupation (observed at the 4-digits level). To this end, I proceed in two steps: I first map each O*NET occupation to its EQF educational requirement Second, I match each O*NET occupation to the corresponding ISCO occupation.

3.3.1. Step 1: matching O*NET occupations to EQF levels

I compute the educational requirements for each O*NET occupation according to the following process:

1. Using the indications given in the EQF, I first define whether a particular skill is provided by each educational level. If it is the case, I define ranges of scores (using the O*NET standardised scores) that are provided by each EQF level. For every skill, I consider that standardised scores lower than 25 in terms of importance and lower than 14 in terms of level are not provided by any educational level[17]. The correspondence between EQF levels and skill scores is reported in Table 1.
2. I then retrieve from the O*NET database the scores of level and importance of each skill in each occupation. As for the level scores, some skills are considered "Not relevant" for some occupations because the related tasks do not require them. I affect a zero value to these occupation-skill couples, leading them not to be linked to any EQF

[15] The *Content* skill category includes Active Listening, Mathematics, Reading Comprehension, Science, Speaking and Writing.
[16] In *Social Skills*, I only keep Instructing and exclude Coordination, Negotiation, Persuasion, Service Orientation and Social Perceptiveness. In *Resource Management Skills*, I only keep Management of Personnel Resources and exclude Management of Financial Resources, Management of Material Resources and Time Management.
[17] Standardised scores lower than 25 in terms of importance and lower than 14 in terms of level correspond to the minimum scores in the respective original rating scales.

level. I also affect a zero value when skills are stated "Not available" in the O*NET data, that is when scores have low precision[18].

3. Using the matches realised in the two previous steps, I define the associated EQF level for each skill in each occupation. For some occupation-skill couples, there are several degrees which can bring the required skill level. In this case, EQF levels are associated with the occupation-skill couple.
4. Finally, I compute the requirements for each skill in each occupation to determine which EQF level brings the larger amount of skills for a particular occupation. At this stage, 768 occupations (88%) are linked to only one particular EQF level, giving the educational requirements directly. For the remaining 105 occupations (12%), several EQF levels (two levels for most of the occupations, three levels for seven occupations) provide the same amount of skills. For these occupations, I keep the two required levels and consider an "upper requirement" and a "lower requirement"[19].

Among the 768 occupations with unambiguous educational requirements, about 35% require EQF 7 or 8 level, 6% require EQF 6 level, 43% require EQF 5 level and 16% require EQF 4 level. None of these occupations are linked to EQF level 3. As for the 105 other occupations, about 80% have two consecutive EQF requirements (mostly EQF 6 and 5 or EQF 7 and 6).

3.3.2. Step 2: matching O*NET occupations to ISCO-08 occupations

The second step of the matching process is to map O*NET occupations into ISCO-08 codes to obtain the EQF requirements for each ISCO occupation. Following Le Barbanchon and Rizzotti (2020), I use the Standard Occupational Classification (SOC) as a pivotal classification. I proceed as follows:

1. The correspondence between ISCO-08 and SOC provided by the International Labour Organization (ILO) uses the SOC 2010 classification, while O*NET is based on the SOC 2018 classification. Using the crosswalk between SOC2010 and SOC2018 provided by the US Bureau of Labor Statistics, I match ISCO-08 to SOC2018. There are 441 ISCO codes for around 850 SOC codes[20], matched "many-to-many": several SOC codes correspond to the same ISCO code, but conversely, in some cases, several ISCO codes correspond to the same SOC code. After the deletion of duplicates (occupations with the same code in ISCO-08 and SOC2018, but not in SOC2010), I obtain a database of 1,147 SOC-ISCO matches.

[18] O*NET recommends suppressing these low precision scores when conducting economic analyses. By affecting a zero value, I ensure that they will not be linked to any EQF level, and thus will not be used in the analysis.
[19] For the seven occupations with three competing EQF levels, I consider the highest of the three as the "upper requirement" and the lowest as the "lower requirement".
[20] More precisely, there are 840 codes in SOC2010 and 867 in SOC2018.

2. The O*NET classification is based on SOC2018, with the six first digits of O*NET codes being exactly the corresponding SOC2018 code. Some O*NET occupations are more precisely defined than in SOC, resulting in two additional numbers. Therefore, one SOC code can correspond to several O*NET occupations, but the opposite does not exist. Using the SOC codes, I match the O*NET occupations to ISCO codes. The matching is once again "many-to-many": several O*NET codes correspond to the same ISCO code, and conversely, several ISCO codes correspond to the same O*NET code. Before any cleaning of the dataset, it contains 1,312 O*NET-ISCO matches.
3. Some ISCO codes do not have any corresponding O*NET code. Many of these codes are armed forces occupations or political occupations, which are generally not included in analyses about overeducation. Thus, I exclude ISCO categories of Armed Forces (ISCO 0) and related occupations (firefighters, police, and prison guards), political positions (ISCO 111), as well as religious occupations (ISCO 2636 and 3413). After these exclusions, 425 ISCO occupations remain, among which 18 do not have any correspondence.
4. Most ISCO codes correspond to several O*NET occupations. However, for 235 ISCO codes, all occupations share the same educational requirements. For 45 additional ISCO codes, all EQF requirements are the same except where requirements are missing[21]. For these ISCO codes (66% of the ISCO occupations), I directly compute the EQF requirement based on the non-missing data.
5. When educational requirements are not the same for all occupations in the same ISCO code (126 ISCO codes), I keep the highest and lowest EQF levels and compute them to the "upper requirement" and the "lower requirement" respectively.
6. For the 18 remaining codes without corresponding O*NET occupations, and thus no EQF level, I affect them the educational requirements of the ISCO occupations sharing the same first three digits (as in step 4) or, if all requirements are not the same, I compute the "upper requirement" and the "lower requirement" (as in step 5).

For about half of the ISCO codes (223), the upper and lower EQF requirements are the same. For the other half, using the upper requirement is likely to lead to an underestimation of overeducation, while using the lower requirement may lead to an overestimation. Indeed, the upper requirement measure is designed not to wrongly consider as overeducated individuals that are actually matched, at the expense of not taking into account a few mismatched workers. On the contrary, the lower requirement measure ensures that all overeducated workers are considered as such, at the cost of considering as overeducated some workers that are actually matched in their job. Using these requirements, the *Skill Analysis* considers individuals as overeducated if their actual EQF level is higher than the (upper or lower) requirement in their

[21] Most codes where requirements are missing correspond in SOC2018 to "All other" or "Not elsewhere classified" categories.

particular ISCO-08 occupation. The upper requirement and lower requirement *Skill Analysis* measures are respectively referred to as "Upper SA" and "Lower SA" in the following[22].

3.4. Advantages and drawbacks of the *Skill Analysis* measure

Using the skills requirements provided by O*NET, the *Skill Analysis* measure associates ISCO-08 occupations to EQF levels. The international dimension of these classifications should allow to use this measure in cross-countries analyses of overeducation (within Europe). Another advantage of this measure is to use occupations rather than occupational groups. Educational requirements are more likely to vary between occupations within the same group than in precisely defined occupations (Halaby, 1994). From this perspective, this approach addresses the aggregation problem, which is one of the biggest issues of overeducation measures (especially statistical measures, for which aggregation of occupations is often required to ensure a sufficient sample size in each occupational group). Therefore, the approach that I propose measures overeducation more accurately than its statistical alternatives.

However, the *Skill Analysis* measure has limitations. First, it links skill levels for occupations and degrees without taking into account professional experience, which can bring and improve skills through on-the-job training, increasing the level of skills of individuals compared to that provided by education. Second, as it relies on the current definition of educational levels, degrees may not correspond exactly to the EQF levels under consideration for individuals who graduated a long time ago. Therefore, the *Skill Analysis* measure is more appropriate for young workers, who are anyway the most affected by overeducation (Baert et al., 2013; Dolton & Vignoles, 2000). Third, for some occupations, several educational levels appear as "normal", raising the question of which one is the most appropriate. This drawback is partially managed by making a difference between a lower and an upper educational requirement for these occupations. Fourth, as for all objective and statistical overeducation measures, it is based on the skills that educational levels should theoretically (according to the EQF) bring to graduates, without information on which skills they actually provide nor on the heterogeneity between degrees at the same level.

[22] The correspondence table between occupations and educational levels according to the *Skill Analysis* measure can be provided upon request.

Table 1 - Learning outcomes expressed in skill scores by EQF level

Skills	EQF 7-8	EQF 6	EQF 5	EQF 4	EQF 3
Active Learning	$I \geq 50$; $L \geq 57$	$I \geq 50$; $57 > L \geq 43$	$50 > I \geq 25$; $57 > L \geq 29$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $29 > L \geq 14$	-
Critical Thinking	$I \geq 25$; $L \geq 57$	$I \geq 25$; $57 > L \geq 43$	$I \geq 25$; $57 > L \geq 43$	$I \geq 25$; $43 > L \geq 14$	-
Learning Strategies	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 29$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 29$	$50 > I \geq 25$; $29 > L \geq 14$	-
Monitoring	$L \geq 43$	$L \geq 43$	$L \geq 43$	$43 > L \geq 14$	-
Instructing	$I \geq 50$; $L \geq 57$	$I \geq 50$; $57 > L \geq 43$	$50 > I \geq 25$; $L \geq 14$ $I \geq 50$; $43 > L \geq 14$	-	-
Equipment Maintenance	-	-	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 29$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $29 > L \geq 14$
Equipment Selection	-	-	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 29$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $29 > L \geq 14$
Installation	-	-	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 29$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $29 > L \geq 14$
Operation and Control	-	-	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 29$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $29 > L \geq 14$
Operations Analysis	$I \geq 50$; $L \geq 57$	$I \geq 50$; $57 > L \geq 29$	$I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 14$	-
Operations Monitoring	$50 > I \geq 25$; $L \geq 57$	$50 > I \geq 25$; $L \geq 43$	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $43 > L \geq 14$	$50 > I \geq 25$; $43 > L \geq 14$
Quality Control Analysis	$I \geq 50$; $L \geq 43$	$I \geq 50$; $L \geq 43$	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 14$ $I \geq 50$; $43 > L \geq 14$	-
Repairing	-	-	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 14$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 14$
Technology Design	$I \geq 50$; $L \geq 43$	$50 > I \geq 25$; $L \geq 14$ $I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 14$ $I \geq 50$; $43 > L \geq 14$	-	-
Troubleshooting	$50 > I \geq 25$; $L \geq 57$	$I \geq 50$; $L \geq 43$	$I \geq 50$; $L \geq 43$	$I \geq 25$; $43 > L \geq 14$	$I \geq 25$; $43 > L \geq 14$
Judgment and Decision Making	$I \geq 50$; $L \geq 43$	$I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 29$	$50 > I \geq 25$; $29 > L \geq 14$	-
Systems Analysis	$I \geq 50$; $L \geq 43$	$I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 29$	$50 > I \geq 25$; $29 > L \geq 14$	-
Systems Evaluation	$I \geq 50$; $L \geq 43$	$I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 29$	$50 > I \geq 25$; $29 > L \geq 14$	-
Management of Personnel Resources	$I \geq 50$; $L \geq 43$	$I \geq 50$; $43 > L \geq 14$	$50 > I \geq 25$; $L \geq 29$	$50 > I \geq 25$; $29 > L \geq 14$	-
Complex Problem Solving	$I \geq 50$; $L \geq 57$	$I \geq 50$; $57 > L \geq 14$	$50 > I \geq 25$; $L \geq 29$	$50 > I \geq 25$; $29 > L \geq 14$	-

Source: Classification based on European Commission (2017), Annex II and O*NET

Note: I denotes the skill score of importance, L denotes the skill score of level. "-" denotes that the skill is not provided by the EQF level.

Lecture: Occupations with a score of importance higher than 50 and a score of level higher than 57 require a degree at EQF levels 7-8 in terms of Active Learning.

4. Empirical comparison of overeducation measures

4.1. Data and statistical measures

In order to explore differences in overeducation measures, I use data from the French Labour Force Survey (LFS) conducted by the French National Institute of Statistics for year 2021[23]. The LFS is a large household survey conducted in European countries and built to ensure comparability between countries, allowing both short-term and structural analyses of the labour market. This dataset provides detailed information on individual socioeconomic (gender, age, educational level, etc.) and job characteristics (wage, occupation, hours worked, etc.). I am then able to compute individuals' EQF levels and link them to the requirements in their particular ISCO-08 occupation.

I exclude Armed forces, Political and religious occupations, as well as self-employed workers. The sample includes 116,512 employed individuals aged 15 to 64 years old, and 20,058 are younger than 30. Among young workers, 48% have a higher education degree: around 16% graduated at each tertiary EQF level (EQF levels 5, 6 and 7-8). The other half of the sample has at most a high school diploma.

Similarly to Capsada-Munsech (2019), I compute three statistical overeducation measures in order to compare them to the *Skill Analysis* measures (SA)[24]. All measures to be compared are summarised and described in Table 2. For statistical measures, I compute overeducation rates at the ISCO-08 2-digits level. Indeed, I am not able to use the 4-digits level due to the small number of observations in each code at this level, not allowing to calculate the statistics accurately. The first statistical measure ("Mode") is based on the mode of the distribution: I consider an individual as overeducated if her or his EQF level is higher than the modal EQF level in her or his particular ISCO occupation. Using ISCO codes at the 2-digits level implies the aggregation of several occupations that are sometimes very different and thus have various educational requirements. Therefore, considering only one level as the "normal" requirement does not allow to take into account the variety of the actual requirements. Instead, I compute a second measure ("2-Mode") that considers as "normal" the two most attained educational levels in the occupation (rather than only one in the previous measure). The third measure ("Q3") is based on the deviation from the median of the distribution. The idea behind this measure is that there is an educational mismatch for the degrees farthest from the median. Using this indicator, I consider that individuals are overeducated if their EQF level is higher than the third quartile of the distribution in their ISCO-08 2-digit code.

[23] At the time of writing, LFS data are not available for more recent years than 2021. Earlier versions of this paper are based on data from LFS 2018 and 2019, which provided similar results.

[24] Unfortunately, the LFS data do not allow me to compare them to subjective approaches, as there is no information about workers' assessment about the level required to do (or to get) their job or their feeling of being overeducated.

Table 2 - Overeducation measures

Measure	*Type*	*Overeducated if…*	*Aggregation*
Upper SA	Objective	educational level above the highest "normal" level as defined in Section 4	ISCO 4-digits
Lower SA	Objective	educational level above the lowest "normal" level as defined in Section 4	ISCO 4-digits
Mode	Statistical	educational level above the most common level in the occupation	ISCO 2-digits
2-Mode	Statistical	educational level above the highest of the two most common levels in the occupation	ISCO 2-digits
Q3	Statistical	educational level above the third quartile of the distribution in the occupation	ISCO 2-digits

Note: SA stands for *Skill Analysis* measure.

4.2. Overeducation incidence across measures

Table 3 shows the differences in overeducation rates depending on the type of measure. As expected, there are large disparities in overeducation rates between measures. For the whole sample, the incidence of overeducation range from 6.86% (upper *Skill Analysis*-SA) to 26.44% (Mode). Unsurprisingly, the upper requirement objective measure presents the lowest rates regardless of the sample. It is due to the fact that it is, by design, the most parsimonious of the measures that I explore. On the contrary, the Mode measure leads to the highest overeducation rates, because of the restrictive definition of the "normal" educational requirements.

Table 3 - Overeducation rates by type of measures

Measure	Upper SA	Lower SA	Mode	2-Mode	Q3
Whole sample *N*=116512	6.86%	17.80%	26.44%	15.40%	10.40%
15-29 year-olds *N*=20058	10.43%	23.39%	41.79%	22.63%	16.55%
Men *N*=55591	5.98%	17.75%	24.73%	13.69%	9.81%
Women *N*=60921	7.67%	17.86%	27.99%	16.97%	10.94%

Source: Author's calculation based on French LFS
Note: SA stands for *Skill Analysis* measure.

Regarding the gender differences in overeducation rates, women are more often mismatched than men, whatever measure is used. However, the gap between men and women is larger using the 2-Mode measure than the alternatives. Given that the 2-Mode measure adds a second "normal" educational level for each occupation, which is for about half of the occupations higher than the first one, this may illustrate the tendency for women to overinvest

in education relative to men, in order to compensate for their higher risk of being discriminated on the labour market (Becker *et al*., 2010). Overall, the 2-Mode measure shows overeducation rates similar to the lower SA (except for men), while the Q3 measure provides lower ones.

Because, by design, the *Skill Analysis* measure is more accurate for young workers (see Section 3.4.), Table 4 displays the overeducation rates for each measure, with detailed rates reported by occupations and EQF levels. Overall, overeducation is higher for young workers, regardless of the measure. The upper SA leads to lower overeducation incidence than statistical measures, while the lower SA provides overeducation rates similar to the ones observed with the 2-Mode. For several ISCO 1 and 2 occupations, I observe overeducation rates equal to zero, especially using statistical indicators. The first plausible explanation is the aggregation problem: while *Skill Analysis* measures are computed using precise occupations, statistical ones are computed on the ISCO 2-digits occupational groups to ensure a sufficient sample size. Thus, the variation in educational requirements between occupations in the same ISCO code may be hidden behind this aggregation, which is not the case for the *Skill Analysis* measures. The other (but not exclusive of the first) explanation is the specific bias of statistical measures, which comes from the design of these measures itself, and can be expressed as "if everybody in an occupation is overeducated, nobody is". In that case, the results reflect hiring standards that are higher than the real need for skills in the occupation. As I focus on young workers, this questions the educational system, which appears not to provide the skills required on the labour market, at least for the concerned occupations.

Regardless of the measure, EQF level 6 graduates are more often overeducated than others. In France, this level corresponds to degrees which are designed to give access to longer studies, but not to enter the labour market directly[25]. Therefore, this finding is in line with Goux and Maurin (1994), who find that years of education without formal certification have lower returns on the labour market than degree-validated educational credentials.

The computation of overeducation rates by ISCO 2-digits codes reveals strong disparities between occupations (Table 4). While some occupations hire less than 5% of overeducated workers, others employ between a third and a half (and, in rare cases, more than a half) of their workforce below their educational level. Most of these occupations with higher overeducation are intermediate professions (associate professionals and clerks), regardless of the measure. As predicted by the *Job Competition Model* (Thurow, 1975) and *assignment models* (Sattinger, 1993), this suggests an overinvestment in education in order to gain access to better positions in more saturated sectors. Indeed, most of the related occupations are in the services, law and administration sectors. This is also in line with Sellami et al. (2020), who find that graduates in Economics, Business and Law are among the most at risk of being overeducated.

[25] EQF level 6 corresponds in France to the degrees *Licence* and *Maîtrise*, which are intermediate steps to Master's degrees. The exception is the *Licence professionnelle*, which has a vocational purpose. However, this particular degree represents less than 5% of the students in 2021 (French Ministry of Higher Education, *MESRI-SIES*).

Table 4 - Overeducation rates by measure - 15-29 year-olds

Measure	Upper SA	Lower SA	Mode	2-Mode	Q3
All occupations and EQF levels (N = 20058)	10.43%	23.39%	41.79%	22.63%	16.55%
All occupations, EQF levels 4 to 8 (N = 15114)	13.84%	31.03%	55.43%	30.02%	21.95%
Men (N = 10339)	8.86%	22.09%	40.52%	21.17%	15.04%
Women (N = 9719)	12.10%	24.76%	43.10%	24.16%	18.13%
Occupations					
12 Administrative and Commercial Managers (N = 137)	0	11.68%	0	0	0
13 Production and Specialized Services Managers (N = 92)	0	1.09%	0	0	0
14 Hospitality, Retail and Other Services Managers (N = 90)	3.33%	6.67%	52.22%	27.78%	27.78%
21 Science and Engineering Professionals (N = 756)	8.73%	19.84%	0	0	0
22 Health Professionals (N = 288)	0	2.43%	0	0	0
23 Teaching Professionals (N = 529)	0.76%	13.99%	71.08%	0	0
24 Business and Administration Professionals (N = 1195)	7.03%	60.25%	0	0	0
25 Information and Communications Technology Professionals (N = 565)	40.53%	76.99%	0	0	0
26 Legal, Social and Cultural Professionals (N = 418)	1.91%	29.90%	0	0	0
31 Science and Engineering Associate Professionals (N = 567)	6.70%	40.56%	32.63%	32.63%	32.63%
32 Health Associate Professionals (N = 823)	4.62%	15.31%	56.62%	1.46%	1.46%
33 Business and Administration Associate Professionals (N = 1239)	9.69%	41.00%	44.71%	44.71%	18.72%
34 Legal, Social, Cultural and Related Associate Professionals (N = 926)	13.61%	22.46%	36.61%	36.61%	23.00%
35 Information and Communications Technicians (N = 180)	10.56%	39.44%	32.78%	32.78%	32.78%
41 General and Keyboard Clerks (N = 608)	35.53%	35.53%	59.21%	59.21%	35.53%
42 Customer Services Clerks (N = 414)	26.09%	29.71%	44.20%	44.20%	28.99%
43 Numerical and Material Recording Clerks (N = 577)	23.92%	38.47%	56.67%	33.10%	33.10%
44 Other Clerical Support Workers (N = 264)	48.11%	50.38%	60.61%	60.61%	48.11%
51 Personal Services Workers (N = 1116)	7.17%	14.78%	62.81%	18.10%	18.10%
52 Sales Workers (N = 1663)	13.11%	21.71%	25.02%	25.02%	25.02%
53 Personal Care Workers (N = 897)	1.90%	5.24%	54.85%	9.03%	9.03%
54 Protective Services Workers (N = 146)	0.68%	10.96%	59.59%	16.44%	16.44%
61 Market-oriented Skilled Agricultural Workers (N = 442)	4.30%	17.42%	55.20%	20.81%	20.81%
62 Market-oriented Skilled Forestry, Fishery and Hunting Workers (N = 38)	10.53%	15.79%	57.89%	15.79%	15.79%
71 Building and Related Trades Workers (excluding Electricians) (N = 850)	3.06%	4.35%	39.53%	6.59%	6.59%
72 Metal, Machinery and Related Trades Workers (N = 841)	1.43%	11.77%	64.21%	21.28%	21.28%
73 Handicraft and Printing Workers (N = 72)	6.94%	12.50%	70.83%	16.67%	16.67%
74 Electrical and Electronic Trades Workers (N = 302)	0.66%	6.95%	69.87%	25.50%	25.50%
75 Food Processing, Woodworking, Garment and Other Craft and Related Trades Workers (N = 553)	12.30%	12.84%	44.67%	18.44%	18.44%
81 Stationary Plant and Machine Operators (N = 455)	2.86%	4.84%	53.85%	12.31%	12.31%
82 Assemblers (N = 134)	2.99%	8.21%	54.48%	16.42%	16.42%
83 Drivers and Mobile Plant Operators (N = 679)	7.81%	9.72%	52.43%	52.43%	10.60%
91 Cleaners and Helpers (N = 500)	7.00%	7.00%	52.80%	52.80%	52.80%
92 Agricultural, Forestry and Fishery Labourers (N = 204)	3.92%	17.16%	57.35%	57.35%	17.16%
93 Labourers in Mining, Construction, Manufacturing and Transport (N = 1118)	13.86%	17.26%	63.15%	17.26%	17.26%
94 Food Preparation Assistants (N = 280)	12.86%	12.86%	60.36%	60.36%	12.86%
96 Refuse Workers and Other Elementary Workers (N = 90)	13.33%	13.33%	50.00%	50.00%	13.33%
EQF levels[a]					
Levels 7-8 (N = 2981)	26.78%	63.35%	39.11%	25.67%	25.67%
Level 6 (N = 3312)	37.71%	72.18%	72.39%	56.89%	45.90%
Level 5 (N = 3289)	8.71%	29.19%	56.69%	50.91%	37.53%
Level 4 (N = 7926)	0	0.08%	54.31%	11.42%	3.43%

Source: Author's calculation based on French LFS

[a] Individuals without any degree (N = 1178) cannot be overeducated. By design, individuals with EQF level 3 (N = 3766) cannot be overeducated either.

By design, the upper SA measure systematically underestimates overeducation and the lower SA measure systematically overestimates it. Therefore, the Q3 measure, which reveals an intermediate incidence of the phenomenon, appears to be more appropriate than the alternative statistical measures. However, while the estimated incidence seems accurate, the Q3 might not consider exactly the right individuals as overeducated, since statistical measures take larger occupational groups into account. This may raise some concerns when investigating the drivers and/or consequences of overeducation. When looking at the overlaps between these measures (Table 5), 10% to 15% of the workers are mismatched in only one of the two dimensions. The correspondence is as expected: crossing the upper SA and the Q3 measures, 8.53% of the workers are statistically but not objectively overeducated, against 2.42% that are objectively but not statistically overeducated. I find an opposite correspondence between the Q3 and the lower SA measures: 3.31% of the workers are overeducated statistically but not objectively, and 10.16% objectively but not statistically. The Q3 measure also presents the highest correlations with the *Skill analysis* measures (Table 6), even though the coefficients remain moderate (between 0.55 and 0.60), which tends to confirm a better, although limited, reliability of the Q3 measure (compared to other statistical measures) as an alternative to objective approaches.

Table 5 - Correspondence between *Skill Analysis* and statistical measures - 15-29 year-olds

	Mode		**2-Mode**		**Q3**	
	Matched	Overeducated	Matched	Overeducated	Matched	Overeducated
Upper SA						
Matched	11279	6677	15090	2866	16246	1710
	(56.26%)	(33.31%)	(75.27%)	(14.30%)	(81.04%)	(8.53%)
Overeducated	391	1701	421	1671	485	1607
	(1.95%)	(8.48%)	(2.10%)	(8.33%)	(2.42%)	(8.02%)
Lower SA						
Matched	10072	5286	13768	1590	14694	664
	(50.24%)	(26.37%)	(68.68%)	(7.93%)	(73.29%)	(3.31%)
Overeducated	1598	3092	1743	2947	2037	2653
	(7.97%)	(15.42%)	(8.69%)	(14.70%)	(10.16%)	(13.23%)

Source: Author's calculation based on French LFS
Note: SA stands for *Skill Analysis* measure.

In line with Capsada-Munsech (2019) and Verhaest and Omey (2006), correlations tend to be higher when considering measures of the same type, *i.e.* between the three statistical measures on the one hand and between the two versions of the objective measure on the other hand, as reported in Table 6. On the contrary, correlations between objective and statistical measures are lower, except for the Q3 measure. This suggests that these different measures do not reveal the same phenomenon, but rather different dimensions of mismatch.

Table 6 - Correlation between overeducation measures

	Upper SA	Lower SA	Mode	2-Mode	Q3
Upper SA	1				
Lower SA	0.6177	1			
Mode	0.2735	0.2704	1		
2-Mode	0.4670	0.5310	0.6383	1	
Q3	0.5536	0.5952	0.5255	0.8233	1

Source: Author's calculation based on French LFS
Note: SA stands for *Skill Analysis* measure.

5. Conclusion

The aim of this work was to develop a new objective measure of overeducation in Europe, based on skills requirements in occupations. Due to the cost of keeping them consistent with the evolutions of the labour market structures, objective measures are missing in several countries, or are based on old occupational dictionaries, which raises concern regarding the accuracy of the measure (van der Meer, 2006). In this respect, I propose an up-to-date measure of overeducation, similar to what O*NET offers for the US but based on the recommendations of the European Commission (2017), where the educational level(s) needed for each occupation is defined based on the skills required to do the job. Moreover, the use of precise occupations rather than larger occupational groups deals with the aggregation problem that arises with the statistical alternatives.

The empirical comparison, using data from the French Labour Force Survey, tends to confirm that different measures are related to different matters. The different approaches only partially overlap: I find that a significant share of workers is statistically but not objectively overeducated. The opposite is also true for some workers, but to a lesser extent. The results also suggest that in the absence of an adequate objective measure a measure based on the median and quartiles of the distribution is a better alternative than other statistical indicators, with higher correlations and overlaps with the *Skill Analysis* measure. Overall, the upper *Skill Analysis* measure lead to a lower overeducation incidence, while the lower *Skill Analysis* is closer to statistical measures.

References

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. https://doi.org/10.1093/oep/53.3.434

Baert, S., Cockx, B., & Verhaest, D. (2013). Overeducation at the start of the career: Stepping stone or trap? *Labour Economics, 25*, 123–140. https://doi.org/10.1016/j.labeco.2013.04.013

Becker, G. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. New York: Columbia University Press.

Becker, G. S., Hubbard, W. H., & Murphy, K. M. (2010). Explaining the worldwide boom in higher education of women. *Journal of Human Capital*, 4(3), 203-241. https://www.jstor.org/stable/10.1086/657914

Capsada-Munsech, Q. (2017). Overeducation: Concept, theories, and empirical evidence. *Sociology Compass, 11*(10), e12518. https://doi.org/10.1111/soc4.12518

Capsada-Munsech, Q. (2019). Measuring overeducation: Incidence, correlation and overlaps across indicators and countries. *Social Indicators Research, 145*(1), 279–301. https://doi.org/10.1007/s11205-019-02112-0

Cedefop. (2018). *Insights into skill shortages and skill mismatch: Learning from Cedefop's European skills and jobs survey*.

Chevalier, A. (2003). Measuring over-education. *Economica, 70*(279), 509–531. https://doi.org/10.1111/1468-0335.t01-1-00296

Croce, G., & Ghignoni, E. (2012). Demand and supply of skilled labour and overeducation in Europe: A country-level analysis. *Comparative Economic Studies, 54*(2), 413–439. https://doi.org/10.1057/ces.2012.12

Davia, M. A., McGuinness, S., & O'Connell, P. J. (2017). Determinants of regional differences in rates of overeducation in Europe. *Social science research, 63*, 67–80. https://doi.org/10.1016/j.ssresearch.2016.09.009

Di Paola, V., & Moullet, S. (2009). Femmes et fonction publique. Un risque calculé de déclassement ? *Travail et emploi*, (120), 47–61. https://doi.org/10.4000/travailemploi.1655

Di Paola, V., & Moullet, S. (2012). Peut-on en finir avec le déclassement ? : Evolution du déclassement dans la fonction publique en début de carrière. *Formation emploi. Revue française de sciences sociales*, (118), 83–101. https://doi.org/10.4000/formationemploi.3610

Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of education review, 19*(2), 179–198. https://doi.org/10.1016/S0272-7757(97)00036-8

European Commission. (2017). Council Recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning. In 2017/C 189/03.

Fleisher, M. S., & Tsacoumis, S. (2012). O* NET analyst occupational skills ratings: Procedures update. *Raleigh, NC: National Center for O*NET Development.*

Forgeot, G., & Gautié, J. (1997). Insertion professionnelle des jeunes et processus de déclassement. *Economie et statistique, 304*(1), 53–74. https://doi.org/10.3406/estat.1997.2556

Ghignoni, E., & Verashchagina, A. (2014). Educational qualifications mismatch in Europe. Is it demand or supply driven ? *Journal of Comparative Economics, 42*(3), 670–692. https://doi.org/10.1016/j.jce.2013.06.006

Goux, D., & Maurin, E. (1994). Education, expérience et salaire. *Économie & Prévision, 116*(5), 155–178. https://doi.org/10.3406/ecop.1994.5704

Groot, W., & Maassen van den Brink, H. (2000). Overeducation in the labour market: A meta-analysis. *Economics of Education Review, 21*(2), 149–158. https://doi.org/10.1016/S0272-7757(99)00057-6

Halaby, C. N. (1994). Overeducation and Skill Mismatch. *Sociology of Education, 67*(1), 47–59. https://doi.org/10.2307/2112749

Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review, 19*(2), 131–147. https://doi.org/10.1016/S0272-7757(99)00050-3

Kucel, A. (2011). Literature survey of the incidence of over-education: A sociological approach. *Revista Española de Investigationes Sociologicas, 1*(134), 125–142.

Le Barbanchon, T., & Rizzotti, N. (2020). The task content of French jobs [Available at SSRN 3653262]. https://doi.org/10.2139/ssrn.3653262

McGowan, M. A., & Andrews, D. (2017). Skills mismatch, productivity and policies: Evidence from the second wave of PIAAC. *OECD Economic Department Working Papers*, (1403). https://doi.org/10.1787/18151973

McGuinness, S. (2006). Overeducation in the labour market. *Journal of economic surveys, 20*(3), 387–418. https://doi.org/10.1111/j.0950-0804.2006.00284.x

Meroni, E., & Vera-Toscano, E. (2017). The persistence of overeducation among recent graduates. *Labour Economics, 48*, 120–143. https://doi.org/10.1016/j.labeco.2017.07.002

Nauze-Fichet, E., & Tomasini, M. (2002). Diplôme et insertion sur le marché du travail : Approches socio-professionnelle et salariale du déclassement suivi d'un commentaire de Saïd Hanchane et Eric Verdier. *Economie et statistique, 354*(1), 21–44. https://doi.org/10.3406/estat.2002.7383

Nordin, M., Persson, I., & Rooth, D. O. (2010). Education-occupation mismatch: Is there an income penalty? *Economics of education review, 29*(6), 1047–1059. https://doi.org/10.1016/j.econedurev.2010.05.005

Robst, J. (2008). Overeducation and college major: Expanding the definition of mismatch between schooling and jobs. *The Manchester School, 76*(4), 349–368. https://doi.org/10.1111/j.1467-9957.2008.01064.x

Rumberger, R. W. (1981). The rising incidence of overeducation in the US labor market. *Economics of Education Review, 1*(3), 293–314. https://doi.org/10.1016/0272-7757(81)90001-7

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature, 31*, 831–880. http://www.jstor.org/stable/2728516

Sellami, S., Verhaest, D., Nonneman, W., & Van Trier, W. (2020). Education as investment, consumption or adapting to social norm: Implications for educational mismatch among graduates. *Education Economics, 28*(1), 26–45. https://doi.org/10.1080/09645292.2019.1680955

Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy, 98*(1), 169–192. https://doi.org/10.1086/261674

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. https://doi.org/10.2307/1882010

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

Tsang, M. C. (1987). The impact of underutilization of education on productivity: A case study of the US Bell companies. *Economics of Education Review, 6*(3), 239–254. https://doi.org/10.1016/0272-7757(87)90003-3

Tsang, M. C., Rumberger, R., & Levin, H. (1991). The impact of surplus schooling on worker productivity. *Industrial Relations, 30*(2), 209–228. https://doi.org/10.1111/j.1468-232X.1991.tb00786.x

van der Meer, P. H. (2006). The validity of two education requirement measures. *Economics of Education Review, 25*(2), 211–219. https://doi.org/10.1016/j.econedurev.2005.02.004

Verdugo, R. R., & Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources*, 629–643. https://doi.org/10.2307/145998

Verhaest, D., & Omey, E. (2006). Discriminating between alternative measures of over-education. *Applied Economics, 38*(18), 2113–2120. https://doi.org/10.1080/00036840500427387

Appendices

Appendix A: Description of EQF levels

Table 7 - EQF levels with corresponding degrees and learning outcomes

EQF level	Degrees	Knowledge, skills and responsibilities
Level 8	PhD	Mobilise the most advanced knowledge, skills and techniques to solve critical and complex problems
		Extend and redefine existing knowledge and professional practice Demonstrate substantial authority, innovation, autonomy and professional integrity
		Develop new ideas or processes at the forefront of work or study contexts
Level 7	Master's degree, Engineering degree, Graduate education (5-year higher education)	Highly specialised knowledge
		Critical awareness of knowledge issues in a field and at the interface between different fields
		Specialised problem-solving skills in order to develop new knowledge and procedures
		Manage and transform work or study contexts that are complex, unpredictable and require new strategic approaches
		Take responsibility for contributing to professional knowledge and practice Review the strategic performance of teams
Level 6	Bachelor's degree, Undergraduate education (3- to 4-year higher education)	Advanced knowledge of a field of work and study Critical understanding of theories and principles
		Advanced skills required to solve complex and unpredictable problems in a specialised field of work or study
		Manage complex activities or projects
		Take responsibility for decision-making in unpredictable contexts Manage professional development of individuals and groups
Level 5	Vocational university degree, Undergraduate education (2-year higher education)	Comprehensive specialised knowledge within a field of work or study, and awareness of the boundaries of this knowledge
		Comprehensive range of cognitive and practical skills required to develop solutions to abstract problems
		Exercise management and supervision in contexts of work or study with unpredictable change
		Review and develop performance of self and others
Level 4	3-year high school degree (general or vocational)	Knowledge in broad contexts within a field of work or study
		Cognitive and practical skills required generate solutions to specific problems in a field work or study
		Self-management within the guidelines of contexts that are usually predictable Supervise the routine work of others
		Take some responsibility for the evaluation and improvement of work and study activities
Level 3	2-year high school degree (vocational)	Knowledge of facts, principles processes and general concepts in a field of work and study
		Cognitive and practical skills to accomplish tasks and solve problems by selecting and applying basic methods and tools
		Take responsibility for completion of tasks
		Adapt own behaviour to circumstances in solving problems

Source: European Commission (2017), Annex II

Appendix B: Description of O*NET skills

Table 8 - O*NET skills

Skills	Description
Process	
Active Learning	Understanding the implications of new information for both current and future problem-solving and decision-making
Critical Thinking	Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions, or approaches to problems
Learning Strategies	Selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things
Monitoring	Monitoring/Assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action
Social Skills	
Instructing	Teaching others how to do something
Technical Skills	
Equipment Maintenance	Performing routine maintenance on equipment and determining when and what kind of maintenance is needed
Equipment Selection	Determining the kind of tools and equipment needed to do a job
Installation	Installing equipment, machines, wiring, or programs to meet specifications
Operation and Control	Controlling operations of equipment or systems
Operations Analysis	Analysing needs and product requirements to create a design
Operations Monitoring	Watching gauges, dials, or other indicators to make sure a machine is working properly
Quality Control Analysis	Conducting tests and inspections of products, services, or processes to evaluate quality or performance
Repairing	Repairing machines or systems using the needed tools
Technology Design	Generating or adapting equipment and technology to serve user needs
Troubleshooting	Determining causes of operating errors and deciding what to do about it
Systems Skills	
Judgement and Decision Making	Considering the relative costs and benefits of potential actions to choose the most appropriate one
Systems Analysis	Determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes
Systems Evaluation	Identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system
Resource Management Skills	
Management of Personnel Resources	Motivating, developing, and directing people as they work, identifying the best people for the job
Complex Problem Solving Skills	
Complex Problem Solving	Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions

Source: O*NET Online Documentation