

# **Factors associated with overeducation among recent graduates during labour market integration. The case of Catalonia (Spain)**

**Joaquin Turmo-Garuz<sup>1</sup>; M-Teresa Bartual-Figueras<sup>2</sup>; Francisco-Javier Sierra-Martinez<sup>3</sup>**

## **Aknowledgement:**

**The authors are grateful to the Institut de Desenvolupament Professional (IDP-ICE) of the University of Barcelona for financial support through the grant REDICE 2016-1200, ACCIO16 AND REDICE18-2180, wich has made this work possible.**

## **Abstract:**

Overeducation has usually been considered a phenomenon brought about during integration into the labour market. There is no single explanation, but rather several factors that are associated with overeducation. We analyse overeducation among graduates in Catalonia four years after finishing their degrees. The analysis is based on the self-assessment made by workers in surveys conducted by AQU (Catalan University Quality Assurance Agency) between 2008 and 2014 and we use logit and probit statistical models to determine the probability of a graduate being overeducated, depending on a wide range of economic, sociological, technological and academic variables. We use the Heckman methodology in the analysis. This study corroborates the results of previous studies on the relationship between wages, job satisfaction and overeducation. In addition, the results show, firstly, differences in levels of overeducation between different fields of study, most notably between 'Humanities and Arts' and 'Health Sciences'. Secondly, the results reveal the impact of the economic cycle on overeducation. Thirdly, the variables used in the statistical model exhibit stable behaviour and, as a result, they provide an explanation for overeducation as a structural phenomenon, regardless of the economic situation. Furthermore, other variables show a significant relationship with signalling theory and career mobility theory, which both explain the overeducation phenomenon. We especially emphasise the role of professional career development during undergraduate studies. Additionally, we found that the family socio-economic environment is relevant in explaining overeducation and, lastly, technological factors and aspects of the graduates' work environment also contribute to explaining the phenomenon.

**Keywords:** Overeducation; university graduates; labour market integration.

---

1 CRIT – Research Group, Department of Economy, Universitat de Barcelona, Av. Diagonal, 696, 08034 Barcelona, Spain. E-mail: jturmo@ub.edu

2 CRIT – Research Group, Department of Economy, Universitat de Barcelona, Av. Diagonal, 696, 08034 Barcelona, Spain. E-mail: bartual@ub.edu

3 Department of Econometrics, Statistics and Applied Economics, Universitat de Barcelona, Av. Diagonal, 690, 08034 Barcelona, Spain. E-mail: fjsierra@ub.edu

## 1. Introduction

Access to education increases the probability of gaining employment and of improving the quality of that employment, as well as the job's productivity and wages (Berg 1970; OECD 2017). Nevertheless, since the publication of Freeman's seminal work *The Overeducated American* (1976), where he detects mismatches between the education level required by a job and the education level of the workers hired, the study of those mismatches has triggered different explanatory hypotheses on their characterisation and their causes.

In general terms, overeducation can be defined as a situation in which a person performs a professional activity whose educational requirements are lower than the academic education received by the person in question. Overeducation is a form of labour market mismatch that has generated numerous studies attempting to analyse its meaning and relevance in the labour market. Authors such as Berg (1970), Freeman (1976) and Smith and Welch (1978) have studied this phenomenon from the very beginning, while McGuinness (2006), García Montalvo (2008), Leuven and Oosterbeek (2011) Capsada-Munsech (2017) and Nieto and Ramos (2017)<sup>4</sup> provide some excellent reviews of the extensive literature on overeducation and its determining factors.

Overeducation can result in negative consequences for the labour market and the production system, such as lower efficiency, lower wages and lower productivity. It is also associated with personal issues, such as lower levels of job satisfaction for employees and less commitment to the firm's activity (Verdugo and Verdugo 1989; Hartog 2000; Iriondo and Pérez-Amaral 2015).

Numerous studies have tried to determine the importance of the phenomenon of overeducation, as well as its relationship with human capital theory and social and economic factors. Thus, authors such as Duncan and Hoffman (1981), Rumberger (1987), Verdugo and Verdugo (1989) or Alba-Ramírez (1993) have analysed the relationship of overeducation with wages, as a return on investment in human capital, and the characteristics of that human capital. Other noteworthy studies include those by Chevalier (2000), Allen and van der Velden (2001), Fleming and Kler (2008), Green and Zhu (2010), Caroleo and Pastore (2013), Diem (2015) and Verhaest and Verhofstadt (2016), among others, who have analysed the relationship between overeducation and job satisfaction. Overeducation throughout professional career development has also been studied (Baert et al., 2013), as has the relationship between field of study and overeducation (Barone and Ortíz, 2011; Caroleo and Pastore, 2013 and 2018). More recently, determinants from the family sociocultural environment have been introduced in studies such as Capsada-Munsech's (2015), and the impact of the economic crisis on the labour market and overeducation has also been explored (Pineda-Herrero et al., 2016; Ermini et al., 2017).

In general, the analysis of overeducation provided in these studies is related to a specific period or cross section; but there are not many analyses of overeducation comparing different periods and different economic circumstances. Furthermore, there are few studies that analyse overeducation from a general or global perspective using the different types of variables found in the statistical model presented here.

The aim of this paper is to analyse the factors associated with overeducation in the process of labour market integration among university graduates in Catalonia in the periods 2004-2008, 2007-2011 and 2010-2014, but with the same set of explanatory variables. This article aims to answer questions arising from the studies mentioned above in order to carry out a broad statistical analysis of overeducation.

- Firstly, whether overeducation is a cyclical or structural phenomenon and, therefore, what stability the variables have over time.
- Secondly, and related to the above, what influence the economic cycle has on overeducation, given that the period studied includes the economic crisis beginning in 2008.
- Thirdly, to what extent the relationship between overeducation and the most relevant labour market theories, such as human capital theory, signalling theory and career mobility theory, can be contrasted. In particular, it is interesting to contrast the importance of the graduates' field of study and the work they carried out before finalising their studies.
- Fourthly, what influence the family's socio-economic environment has on overeducation.
- Fifthly, what role the economic and technological factors of the work environment play.

---

<sup>4</sup> Some authors such as Allen and van der Velden (2001) and Nieto and Ramos (2017) distinguish between overeducated workers, who have acquired formal education, and overskilled workers, who have acquired skills additional to their academic education. In this paper, we consider only overeducation.

- Lastly, whether there are gender differences in overeducation.

Given that only a part of the survey cases was selected (people working at the time of responding to the survey), a two-step logit approach was used, following the Heckman method (1979). The estimations made allow us to confirm most of the hypotheses, providing interesting results both in terms of the stability of the variables in the model and their explanatory power in relation to the phenomenon. The results are relevant for most of the factors and, in general, are in agreement with those obtained in the studies cited. However, the significance of some variables and, consequently, their explanatory power are noticeably altered as a result of the economic crisis. Furthermore, as well as observing the role of those variables, the aim is to obtain some results with policy implications.

As we shall see, some of the limitations of this study include the subjective nature of measuring overeducation and the determination of the variables and their relationship with the theoretical principals.

The following section of the paper presents a review of labour market theory and of the factors related to overeducation. Section 3 presents the criteria for measuring overeducation and section 4 establishes the analysis methodology. In section 5, the statistical model is introduced and the results are analysed and discussed and, lastly, the conclusions are put forward in section 6.

## **2. Theoretical framework: Overeducation, the labour market and related factors**

Overeducation is a phenomenon whose conceptualisation has generated controversy and it has not made an easy theoretical fit in labour market analysis. The studies carried out since Freeman (1976) do not so much aim to develop a theory of overeducation as to try to find an explanation for it, as well as its economic and social consequences, based on labour market theory.

### **2.1. Overeducation, the labour market and related factors**

There are several studies on overeducation. One set of studies relates overeducation to human capital theory and other theories on workers' recruitment or employability, other studies emphasize on factors associated with overeducation. Economic cycle, technical change and field of study are also considered.

In the first set of studies, according to Becker (1964) and Mincer (1974), the human capital accumulated by people during their education and their employment experience prepares them to access jobs where they can develop a level of productivity that, at the same time, should be on a par with their wages. Moreover, there should also be financial balance between those wages and the investment previously made in human capital by the people employed, their wages being the return on their investment. Following this, there would be no room for overeducation in employment, especially in the case of graduates.

As we mentioned before, analysis from the point of view of human capital theory has generated empirical studies by authors such as Duncan and Hoffman (1981), Rumberger (1987), Verdugo and Verdugo (1989), Alba-Ramirez (1993), Hartog (2000) and Rubb (2003), who have observed a lack of balance between prior investment in human capital and the wages received subsequently by employees. A negative correlation has also been observed between workers' wages and overeducation, meaning that there are people with a higher education level than that required by their jobs; those people are overeducated for the posts they hold.

Deeper analysis of workers' labour market integration processes has generated subsequent theories and empirical studies. In general, the theories have been developed based on human capital theory and they tend to explain overeducation on the basis of its temporary nature and the lack of information between employees and employers in the labour market integration process. Empirical studies have often emphasised the circumstances and the socio-economic environment associated with overeducation and have focused on detecting the phenomenon through the factors to which it is related. Both types of studies have been the source of many variables used in this statistical analysis.

Credentialist theory considers education level as a filter that classifies individuals in relation to professional tasks; therefore, education level is the point of reference when hiring workers (Arrow, 1973). In this regard, Thurow (1975) and Duncan and Hoffman (1981) consider education level as a signal in recruitment. That is, employers hire on the basis of education level or of the signal emitted by the worker, since that signal or education level reveals the cost of training employees: the higher the education level, the lower the training cost. In this context, the people hired compete for the post; therefore, an increase in labour supply can cause overeducation.

In signalling theory, Spence (1973) considers that there is information asymmetry between employers and employees when it comes to recruitment. Indeed, while workers know their degree of skill and knowledge, employers, in theory, have no way of verifying that information and therefore go by the observable signs or 'signals' of the market: education level. For that reason, in order to cover themselves against the lack of information on hireable individuals, firms will tend to hire those that stand out in terms of the signal emitted, regardless of the education level required to carry out the post's functions. This behaviour can result, at least at the beginning of employment activity, in overeducation, especially if there is an overabundance of graduates in the market.

More recent contributions to these two theories consider relevant elements of job match in recruitment to include proper identification of the credentials and signals the graduates provide (Tan, 2014; Di Stasio et al., 2016). Indeed, proper identification of the credentials provided by graduates, as well as their suitability to the job, can lead to a good match and, as a result, less overeducation.

Career mobility theory also assumes the existence of overeducation at the start of people's employment activity, although, over time, changing firm or promotion within the same firm could allow them to overcome this mismatch (Sicherman and Galor, 1990). Therefore, these authors associate overeducation with graduates' professional career development, such that the characteristics of their training and their professional strategy can influence the success of their labour market integration; that is to say, its speed and quality. A less-analysed aspect in the study of overeducation is the experience of part-time professional work during the period of study. However, Passaretta and Triventi (2015) obtained an interesting result from a practical point of view when they observed a negative relationship between work carried out during undergraduate studies and the risk of subsequent overeducation. A broad perspective on the concept of career mobility should include this aspect in its definition prior to recruitment.

Among the second type of studies, which focus on the factors associated with overeducation, the first thing worth highlighting is job satisfaction. Indeed, a negative relationship has been confirmed between overeducation and levels of job satisfaction (Allen and van der Velden 2001; Verhaest and Omey 2009; Kucel and Vilalta-Bufi 2013; Diem 2015). This is consistent with the negative relationship between wage level and overeducation, as mentioned above. As a result, there is often a correspondence between increased overeducation, lower earnings and lower job satisfaction. The explanation is based on the fact that the underutilisation of the employee's knowledge results in lower efficiency, lower productivity and lower wages (Caroleo and Pastore 2013; Diem 2015). As a result of this, overeducation can give rise to lower job satisfaction and, consequently, lower motivation, involvement in and commitment to the firm's activities and, specifically, decreased participation in internal promotion (Hersh 1991). Furthermore, an overeducated person's continuance in his or her post increases the risk of cognitive obsolescence, since by not utilising the knowledge and skills acquired they are neither consolidated nor updated (Baert et al. 2013). From another perspective, overeducation can also give rise to increased instability in the post: overeducated people have a greater incentive to look for another job where they could develop their professional career (Sicherman and Galor 1991).

Authors such as Capsada-Munsech (2015), Gaeta (2015) and Erdsiek (2016) have considered the family sociocultural and economic context as a relevant factor in the design of graduates' professional career strategies. Elements that are unrelated to their academic training but that form part of their environment have been considered to be related to the probability of overeducation. The introduction of these elements is justified by the fact that the most immediate environment is the context in which graduates' main decisions about their studies and employment future develop; in other words, it is the context in which they obtain and process most of the information that will determine their decisions. This involves indirectly gathering data on aspects such as the social capital found in the environment, as well as the person's ability to obtain information about the labour market (Verhaest and Omey 2010). Thus, part of an employed person's probability of overeducation can be explained by that immediate environment. In general, the authors cited have obtained results that are consistent with what might be expected: a higher sociocultural level is inversely related to the probability of overeducation.

The relationship between the economic cycle and overeducation has been studied by authors such as Pineda-Herrero et al. (2016), Ermini et al. (2017) and Summerfield and Theodossian (2017), who all found a negative relationship: an increase in overeducation during the period of economic recession. In our case, the phenomenon can be directly observed in Table 3, where the percentages of overeducation in 2014 (reflecting labour market integration since 2010) are noticeably higher than those in 2008 and 2011. However, the aim is also to analyse the possible impact on the behaviour of the explanatory variables.

Technical change can also reduce overeducation among the most skilled workers. Indeed, Katz and Murphy (1992) have observed that some advanced countries have experienced a considerable rise in the demand for

skilled labour, at the same time as the wage rate of highly skilled workers has increased in relation to the average wage rate. According to Acemoglu (1998 and 2002) and Autor et al. (2003), these changes in the labour market and in wages are explained by there having been a skill-biased technical change (SBTC); that is, the greater growth of the demand for skilled labour over its supply has enabled not only the absorption of most graduates emerging from universities, but also growth in the relative wages of the highest paid. This would have mitigated the risk of the emergence of overeducation. Thus, Di Pietro (2002) and Ghignoni and Verashchagina (2014) consider that the differences in the speed of technical change and its impact on the production apparatus could help explain differences in the levels of overeducation between countries, such that countries with greater technological investment and greater innovation tend to make better use of the skilled workforce and, consequently, they tend to have a lower level of overeducation. Within each country, there are technological differences between sectors or firms that, in practice, result in differences in the speed of adaptation of each sector or firm to technical change. Therefore, slow adaptation to technical change means a low technological level and lower demand for highly skilled workers and, consequently, there will be a higher probability of overeducation; on the other hand, if adaptation is fast, the complete opposite will occur (Rumberger 1987; Hartog and Oosterbeek 1988; Dolton and Vignoles 2000; McGuinness 2006).

Lastly, the graduate's field of study has some relationship with overeducation. Caroleo and Pastore (2013 and 2017) and Capsada-Munsech (2015) have studied this relationship in the case of Italy, noting differences based on field of study; Verhaest et al. (2017) have also studied it in several different countries. The causality of this relationship is complex and derives from the productive structure of the economies. However, insofar as it defines the graduate's area of knowledge, it can contribute to the signalling mentioned above.

Therefore, on the basis of the theoretical review, we can identify the main general factors associated with overeducation:

- those related to labour market integration, such as wages or job satisfaction;
- signals or credentials, as well as academic performance;
- career mobility before and after recruitment, such as activities carried out during undergraduate studies and seniority;
- family economic and sociocultural level;
- aspects related to the economic and technological conditions of the work environment.

We can define the variables that may have a probabilistic relationship with overeducation on the basis of these general factors, following the model presented below.

## 2.2. Measuring overeducation

Measuring overeducation comes with the difficulties pertaining to empirical work and those derived from the choice of measurement approach and from the definition of the concept<sup>5</sup>. There are three approaches to measuring overeducation: worker self-assessment, statistical analysis and analysis of the job (Chevalier 2003; Verhaest and Omeij 2006; García Montalvo and Peiró 2009; Leuven and Oosterbeek 2011).

*Self-assessment* is based on workers' responses regarding the suitability of their education level to the requirements of their job, whether through direct or indirect questions (Verhaest and Omeij 2006). The *statistical approach* determines the match between the education received and the 'most frequent' education level in the job in question; that is, according to the education level of the people that normally hold that post (Verdugo and Verdugo 1989). Hence, individuals with an education level that is more than one standard deviation above the mean education level within their occupation would be overeducated. *Analysis of the job* takes the qualitative classification of jobs following the ISCED classification as its reference; this way, workers that, according to the standards, show a higher education level than that corresponding to the job in question are overeducated.

---

<sup>5</sup> In fact, some authors (Verhaest et al., 2017) distinguish between vertical and horizontal overeducation, or what in Table 3 are referred to as overeducation and mismatch, respectively. In this article, we have used the latter terminology, which could be seen as restrictive, but which allows us to better delimit the term and its meaning with regard to verticality or hierarchy in the conditions of labour market integration.

Verhaest and Omeij (2006) and Leuven and Oosterbeek (2011) carried out critical analyses of the different approaches. According to these authors, the statistical approach is the least desirable, mainly due to the instability of the standard deviation of education levels in relation to the mean, resulting from variations in labour supply and demand that can affect the average education level used for each post. Together with Hartog (2000), these authors consider that the ideal approach would be the analysis of the job, as long as the standards of the job functions are continually updated.

Self-assessment raises the issue that subjectivity can prompt the workers interviewed to overestimate the requirements of the post, which can in turn lead to an underestimation of overeducation (Hartog 2000). However, given that in this case we are dealing with an analysis carried out at three different times, only the graduates' own assessment of the suitability of their education to the post can provide greater understanding of how the conditions for job match in the labour market are updated over time. Furthermore, the availability of data that allow us to make an immediate and reliable calculation of overeducation is another factor that supports the selection of the self-assessment approach.

### 3. The empirical analysis and the model

The empirical analysis was carried out using the results of the surveys conducted by AQU (Catalan University Quality Assurance Agency) with university graduates every three years. The surveys used were from 2008, 2011 and 2014, and were conducted with graduates in 2004, 2007 and 2010, respectively (Table 1)<sup>6</sup>. The population was composed of 59.5 % women in 2008, 59.1 % in 2011 and 59.8 % in 2014; these percentages are similar to the make-up of the samples in terms of sex (61.2 %, 60 % and 59.1 % for each survey). The size of the total sample varies between 30 % and 50 % of the graduate population in Catalonia. The study sample includes only graduates that entered the labour market after their graduation and with a maximum of four years' professional experience. The result is a sample that varies between 32 % and 40 % of the size of the population<sup>7</sup>. The size of the sample in relation to the population provides highly consistent results; in addition, the missing data show a random pattern without any bias.

Table 1. Characteristics of the surveys and populations

| Survey | Population | Total sample | %    | Sampling error | Study sample | %    | Sampling error |
|--------|------------|--------------|------|----------------|--------------|------|----------------|
| 2008   | 24 169     | 12 553       | 51.9 | 0.62           | 9 660        | 40   | 0.79           |
| 2011   | 28 616     | 16 182       | 56.5 | 0.52           | 10 933       | 38.2 | 0.75           |
| 2014   | 31 279     | 17 337       | 55.4 | 0.51           | 10 263       | 32.8 | 0.81           |
|        |            |              |      |                |              |      |                |

Source: compiled by authors based on data from AQU.

The surveys contain between 75 and 82 questions providing information on the academic aspects of each graduate, their work environment and their integration process, and data on their personal skills and their personal environment.

In this paper, overeducation is determined using the graduates' responses, first, to the objective question (a) of whether a university degree is required or not and if it is specific and, second, to the more subjective question (b) of whether the tasks carried out correspond to the academic training received by the graduate. Table 2 summarises the cases. The responses come down to a dichotomy between normal labour market integration and overeducation among graduates in their posts.

<sup>6</sup> Information about the graduates' university of origin and their race/ethnicity was excluded from the surveys for legal reasons. That meant a significant loss of information.

<sup>7</sup> The increase in unemployment during the period of economic crisis caused a reduction in the number of valid responses from employed graduates.

Table 2. Assessment of graduates' responses

| Graduate responses         |                               | Assessment according to responses |                                  |
|----------------------------|-------------------------------|-----------------------------------|----------------------------------|
| Qualification required     | Activity carried out          | Match/Mismatch                    | Normal integration/Overeducation |
| Specific university degree | Suitable to degree            | Match                             | Normal integration               |
|                            | Not suitable to degree        | Mismatch                          |                                  |
| General university degree  | Suitable to university degree | Match                             |                                  |
|                            | Not suitable                  | Mismatch                          | Overeducation                    |
| No degree                  | Suitable to university degree | Match                             | Normal integration               |
|                            | Not suitable                  | Mismatch                          | Overeducation                    |

Source: AQU and compiled by authors.

The statistical analysis was done by applying a binary logit regression model that estimates the probability of being overeducated or not. The variable to be explained takes the value of 0 for people that are in a 'normal' work activity (i.e. suited to their level of training) and the value of 1 for people that are overeducated in their post. The equation to be estimated determines the probability of a graduate being overeducated, according to the following formula:

$$P(Y = 1) = 1/(1+e^{-(Z_i)});$$

P being the probability of overeducation and

$Z_i = \beta_1 + \beta_2 + \dots$  the equation for the set of explanatory variables of overeducation, in which  $\beta_i$  are the variables' coefficients.

The model was adapted to the conditions observed in the variables, particularly to the level of overeducation in each period analysed, following the data in table 3. Thus, the cut-off points for classifying cases have been defined according to the percentages of overeducation observed in each period. As a result, the model's parameters are the average values for each period: 0.124 for the estimation corresponding to 2008, 0.118 for 2011 and 0.176 for 2014.

There were 24 variables selected (seven qualitative and 17 quantitative), which, with the relevant transformations, cover the factors related to overeducation mentioned above (see end of section 2.1).

Principal component analysis with orthogonal rotation (varimax) was performed on 12 variables, which meant obtaining values that were independent of each other by grouping the original variables and reducing them to four factors. The varimax method can provide non-correlated factors. As a result, the correlation between factors and variables can be observed more clearly than with other methods of analysis, such as oblimin, oblique rotation or confirmatory factor analysis. Factor reduction was carried out on the variables satisfaction with job content, prospects for improvement, wage level and general work environment (satisf1, satisf2, satisf3 and satisf5), as well as the variables covering skills in the form of theoretical knowledge, practical knowledge, command of languages, command of ICTs, ability to manage and plan and capacity for teamwork (empleab1, empleab2, empleab3, empleab4, empleab5, empleab6 and empleab7). The result gave us four variables that encapsulate these aspects (see Annex 1). The new variables were chosen following what Hair et al. (1999) call the 'selection of surrogate variables'; that is, we selected the four variables that summarise the original set of 11 variables. These variables are the most representative for each of the factors extracted and have a greater factor loading, indicating that they match the corresponding factor:

- INST\_PERS\_COMP: Instrumental personal skills, including social skills, ability to manage and plan and capacity for teamwork;
- JOB\_SATISF: Satisfaction with job content, prospects for improvement, wage level and overall satisfaction;
- THEO\_PRAC\_KN: Degree of theoretical and practical knowledge;
- INST\_KN: Instrumental knowledge in languages and ICTs.

Thus, the 17 variables used are defined as follows (see Annex 3).

*Variables related to labour market conditions:*

- *wages*: gross annual income (EARNINGS) was included in the model as a quantitative variable with eight income intervals ranging from less than €9 000 per year to over €40 000 per year, following data taken from the surveys;
- *degree of job satisfaction*: overall satisfaction with the content of the job (JOB\_SATISF) obtained through principal component factoring, reflecting different aspects of job satisfaction (see Annex 2).

We also used *variables related to academic signals and education characteristics*. It must be taken into account that the number of years of schooling is the same for all of the graduates, which means similar human capital endowment, except for graduates in medicine. However, what do differ between individuals are the characteristics of the individuals' training, which make up that human capital endowment and determine its suitability to the job. The variables selected are:

- MARK, the mark awarded on graduating, a continuous quantitative variable with values between 0, minimum, and 5, maximum, and shows the average mark of the graduate studies;
- GLOB\_EDU, the overall level of university training, an ordinal qualitative Likert-scale variable with values ranging between 1 and 7;
- THEO\_PRAC\_KN, the level of theoretical and practical training reached, a variable obtained by principal component factoring using the aspects corresponding to both types of training during the graduate's studies (see Annex 1);
- KN\_THEO, the adaptation of theoretical knowledge to the job, a variable obtained directly from the questions the graduates were asked, with values between 1 and 7, in Likert format.

Other variables that are less visible in recruitment were also introduced. They correspond to *personal skills* acquired during the training period. These variables are the result of principal component factoring, covering both aspects as instrumental elements in graduate training (see Annex 1):

- INST\_KN, command of languages and information and communication technologies;
- INST\_PERS\_COMP, the capacity for management, relationships and team work.

*Variables related to career development:*

- SENIORITY is the variable that measures career development within the firm, following Sicherman and Galor's (1991) hypothesis. It is a discrete quantitative variable that oscillates between 0 and 4 years.
- EST&WORK indicates whether there was employment activity during the two final years of the graduates' studies and if it was related to their studies. This variable indicates employment activity prior to recruitment by the firm. It is a polytomous nominal qualitative variable with three values describing the activity carried out by graduates during their studies: they only studied (EST&WORK1); they studied and worked in an area related to their studies (EST&WORK2); and they studied and worked in an area unrelated to their studies (EST&WORK3).

Variable indicating the *sociocultural level* of the graduates' immediate environment:

- LEVEST\_P is the variable used to measure the level of studies of the graduates' parents. According to Verhaest and Omey (2010), the family environment can more directly influence graduates' decision-making regarding their careers. This is an ordinal qualitative variable with five values classifying the education level of the graduates' parents: level 1 corresponds to cases in which both have a primary education; level 2 corresponds to cases when at least one of the parents has a secondary education; level 3 corresponds to cases when both have a secondary education; level 4 is for when at least one of the parents has a university education; and, lastly, level 5 is for when both have a university education.

The *working conditions* refer to the type of working day (TIPJORN) and the type of contract (TIP\_CONT). Lastly, the variables type of contract and type of working day are both dichotomous qualitative variables, which take the values of 0 and 1; in the first case, for a permanent contract and a temporary contract; in the second case, for a full-time working day and a part-time working day.

We can distinguish between the *structural factors* of overeducation from the supply side perspective of the labour market (FIELD\_STUDY) and from the demand side perspective (SECTOR\_7 and NUM\_PER). These variables therefore link overeducation to the production system.



The *variable field of study*, FIELD\_STUDY, is a polytomous nominal qualitative variable that defines the graduate's field of specialisation. This variable has seven values corresponding to the academic areas.<sup>8</sup> It allows us to evaluate the suitability of the skilled jobs supply to the production system, depending on the match of each of the academic areas.

The *structural variables* included in the model are the economic sector (SECTOR\_7) and the size of the firm, that is, the number of employees in the firm (NUM\_PER). The 26 sectors into which economic activity was classified in the surveys were also reduced to seven (as observed in Annex 3), expressed in a polytomous nominal qualitative variable with those values. In terms of the number of workers, the firms were classified into six types according to their size.

Following the OECD (2006), the sectors 'ICT, Finance and Business Services' and 'Health and Education' were considered the most knowledge-intensive and, therefore, to have the highest demand for skilled workers, together with the 'Construction' sector, in terms of its demand for graduates<sup>9</sup>. As a result, the growth of these sectors can contribute to reducing the probability of overeducation among their employees. On the other hand, the role of the rest of the variables is less predictable given the differences between firms, even of the same size and from the same sector, and the variety of contract types.

In addition to all of the variables defined above, the variable SEX was included to examine the effect of gender on overeducation.

Lastly, we considered including variables related to the economic cycle, such as unemployment by field of study and by economic sector, as well as the average growth rate for each period, but they were not significant (see section 4.1).

Having self-selected only graduates who were working at the time of the survey, there was a risk of sample selection bias. Heckman's two-step method (Heckman, 1979) was used to avoid this risk. This method consists of first estimating a probit model to calculate the probability of an individual deciding to work or not. Although this method is used prior to the OLS estimations for wage behaviour (Dolton and Vignoles, 2000; Vatta et al., 2016; Thenuwara and Morgan, 2016; and Caroleo and Pastore, 2018), the study's logit model was adapted using the same statistical basis. For this purpose, we selected variables from the sample that have an impact on the individuals' employability. These are variables relating to the individuals' decision to work or continue studying and signalling variables:

- the variable that reflects the decision to continue studying (*contestu*) has five options: 1) specialised courses; 2) another degree; 3) post-graduate degree or master's; 4) PhD; and 5) other;
- the signalling variables are MARK, which was already introduced in the model, plus four variables which were not introduced in the model and which indicate suitability for recruitment: *aequip* (team work suitability), *apra* (practical training suitability), *ateor* (theoretical training suitability), *asolprob* (problem-solving ability suitability).

The Mills ratio was obtained from the estimation and its inverse was incorporated into the model as another regressor. The significance of this coefficient indicates the extent of the bias that would have been caused if the Mills ratio had not been incorporated into the target model.

Inverse Mills ratio (IMR):

$$\lambda(Z) = \phi(Z)/\Phi(Z)$$

IMR [ $\lambda(Z)$ ] is a new regressor in the target equation.

#### 4. Results and discussion

The results obtained show consistency in the relationship between the explanatory variables and the explained variable. We would also note that the goal is to estimate the probability of vertical overeducation and we did not, therefore, take into consideration the mismatch contained in Table 3 (see footnote 2).

The two-step estimation, introducing Heckman's IMR into the subsequent logit model, shows that the IMR parameter is not significant in 2008 or 2011, but it is in 2014. The implementation of that methodology is

<sup>8</sup> In order to be able to include this variable in the model, the 37 degrees that appear in the survey had to be reduced to the seven areas that appear in the summary of the variables in Table 3, for FIELD\_STUDY.

<sup>9</sup> In the case of the 'Construction' sector, its demand for graduates is mainly in the technical areas of architecture and engineering.

therefore relevant, even if the results have only experienced slight modifications with regard to the estimations made without it.

Nevertheless, it should be noted that the subjective nature of measuring overeducation and the determination of the variables used are the main limitations of the analysis carried out, particularly in terms of the possible underestimation of overeducation (Hartog, 2000). In footnote 3, we mention two more limitations regarding the data: the exclusion of the graduates' university of origin and of the racial/ethnic make-up of the sample. In addition, we must take into account other limitations in the analysis caused by the number of sectoral classifications and university degrees used (which in both cases was too low) due to the restrictions of the model. Lastly, we have not been able to contrast the impact of the higher education reform process (Bologna Process) because it began to be implemented gradually in Spain as of 2008 for students who began their studies in succeeding years. Nevertheless, in Spain, the degree structure remained the same both before and after the Bologna reform: four years for all degrees.

#### 4.1. Evolution of overeducation by academic areas during the crisis

Table 3 shows the results (in percentages) of labour market integration for all of the employed graduates from the different academic areas in Catalonia in the three surveys analysed (2008, 2011 and 2014), as well as the unemployment rates for each field of study. Given that Catalonia is an important Spanish region (autonomous community), the data on graduate unemployment and economic growth are similar to those of other important communities, such as Madrid, and, consequently, to the overall Spanish economy. Indeed, the average rates of economic growth during this period were similar: in the case of Catalonia, 2.9 % during the period 2004-2008, - 0.3 % during the period 2007-2011 and - 0.9 % during the period 2010-2014; while the average rates for the Spanish economy were 3.2 %, 0.0 % and - 0.8 %, respectively. Therefore, the economic crisis equally affected Catalonia and the rest of Spain, although neither the rates of economic growth nor the unemployment rates by field of study or by sector were significant in any of the three periods, therefore demonstrating the greater import of structural factors over cyclical ones in overeducation.

The results of the overeducation analysis show that during the period of economic crisis, job match has deteriorated in the labour market and there has been an increase in overeducation in all academic areas. These results are similar to those of authors such as Pineda-Herrero et al. (2016) or Ermini et al. (2017).

Moreover, the results differ among graduates according to the field of study, with those from the fields of 'Humanities and Arts' having a higher rate of overeducation, while 'Health Sciences' and 'Technology' graduates have the lowest rate. This result, which is similar to that of studies such as Barone and Ortiz (2011) or Caroleo and Pastore (2013 and 2018), reveals a labour market bias that provides advantages for scientific and technical activities over humanistic ones and which has worsened with the economic crisis. In addition, it reveals that sectors requiring the highest levels of scientific and technological knowledge are the ones that perform best during the crisis (Pareja-Eastaway and Turmo-Garuz, 2013; Reig, 2017). The increase in job mismatch and in overeducation during the period of economic crisis is probably associated with the economic system's difficulties with absorbing graduates during that period. The crisis has affected each academic area differently. For example, in the 'Technology' area there was a considerable increase in overeducation, although its level in 2014 was still lower than the average. The same goes for the areas 'Education and Communication' and 'Experimental Sciences and Mathematics'.

Table 3. Job match, overeducation and unemployment from 2008 to 2014 (in percentages)

| Field of study              | 2008  |          |         |              | 2011  |          |         |              | 2014  |          |         |              |
|-----------------------------|-------|----------|---------|--------------|-------|----------|---------|--------------|-------|----------|---------|--------------|
|                             | Match | Mismatch | Overed. | Unemployment | Match | Mismatch | Overed. | Unemployment | Match | Mismatch | Overed. | Unemployment |
| Humanities and Arts         | 61.6  | 6.5      | 31.8    | 7.9          | 63.0  | 4.7      | 32.2    | 16.3         | 47.0  | 11.2     | 41.8    | 21.5         |
| Philology                   | 75.8  | 6.3      | 17.9    | 4.6          | 79.2  | 5.5      | 15.4    | 11.2         | 70.1  | 9.5      | 20.4    | 9.8          |
| Social Sciences             | 78.3  | 4.6      | 17.1    | 3.3          | 78.8  | 4.3      | 16.9    | 9.0          | 71.8  | 7.1      | 21.1    | 11.2         |
| Education and Communication | 87.3  | 4.2      | 8.5     | 2.7          | 88.8  | 4.1      | 7.1     | 6.2          | 75.1  | 6.1      | 18.8    | 11.0         |
| Sciences and Maths          | 79.9  | 7.7      | 12.4    | 2.8          | 82.0  | 6.7      | 11.3    | 8.9          | 78.9  | 5.0      | 16.1    | 12.6         |
| Health                      | 94.9  | 1.5      | 3.6     | 2.4          | 94.9  | 2.0      | 3.1     | 5.4          | 93.1  | 1.6      | 5.3     | 8.0          |

|            |      |     |      |     |      |     |      |     |      |     |      |      |
|------------|------|-----|------|-----|------|-----|------|-----|------|-----|------|------|
| Technology | 84.7 | 7.1 | 8.2  | 2.4 | 80.9 | 7.3 | 11.7 | 7.8 | 77.5 | 6.7 | 15.9 | 7.8  |
| Total      | 82.4 | 5.2 | 12.4 | 3.2 | 83.2 | 5.0 | 11.8 | 8.1 | 76.3 | 6.1 | 17.6 | 10.5 |

Source: compiled by authors based on data from AQU 2008, 2011 and 2014.

#### 4.2 Factors related to overeducation.

Table 4 shows the results of the probit estimations from the first Heckman step for each period, from which the Mills ratio was obtained. The inverse Mills ratio (IMR) was then incorporated into the target equations for each period. Tables 4, 5 and 6 show the results of the logit estimations of the probability of overeducation for each period carried out in two ways: without the incorporation of the IMR (Heckman two step method) into the target equation and with the incorporation of the IMR. In addition, the absence of collinearity between the variables of the equations for the three periods has been confirmed. Both the degree of significance and the classification potential are sufficiently high for this type of model.<sup>10</sup>

In general terms, the results of the second equations (Heckman) show a stable relationship between the explanatory variables and the variable explained in the three periods, although some small differences linked to the economic cycle appear in the last period (2010-2014). Therefore, the probability of a graduate being overeducated four years after finishing his or her studies depends, basically, on the same factors in both 2008 and 2014; consequently, overeducation appears as a structural phenomenon. The differences between variables in the last period suggest that the economic cycle can have a negative impact on overeducation, but with the structural aspect being stronger than the cyclical one. This is one of the study's most interesting findings.

Table 4. Results of the model's explanatory variables in the first period

|                | Logit binary 2004-2008 |               |              | Heckman model 2004-2008 |               |              |
|----------------|------------------------|---------------|--------------|-------------------------|---------------|--------------|
|                | B(coefficient)         | Standar error | Significance | B(coefficient)          | Standar error | Significance |
| SENIORITY      | -.0730                 | .037          | .048         | -.0710                  | .037          | .054         |
| EARNINGS       |                        |               | .000         |                         |               | .000         |
| EARNINGS(1)    | 3.122                  | .441          | .000         | 3.114                   | .441          | .000         |
| EARNINGS(2)    | 2.919                  | .414          | .000         | 2.918                   | .414          | .000         |
| EARNINGS(3)    | 2.459                  | .400          | .000         | 2.457                   | .401          | .000         |
| EARNINGS(4)    | 2.233                  | .394          | .000         | 2.234                   | .394          | .000         |
| EARNINGS(5)    | 1.464                  | .385          | .000         | 1.466                   | .385          | .000         |
| EARNINGS(6)    | .849                   | .390          | .029         | .849                    | .390          | .029         |
| EARNINGS(7)    | .699                   | .405          | .084         | .704                    | .405          | .082         |
| EST&WORK       |                        |               | .000         |                         |               | .000         |
| EST&WORK(1)    | -.071                  | .100          | .477         | -.069                   | .100          | .490         |
| EST&WORK(2)    | -.439                  | .113          | .000         | -.438                   | .113          | .000         |
| GLOB EDU       | -.271                  | .032          | .000         | -.267                   | .032          | .000         |
| INST KN        | .103                   | .046          | .025         | .105                    | .046          | .021         |
| INST PERS COMP | .202                   | .048          | .000         | .213                    | .049          | .000         |
| JOB SATISF     | -.234                  | .042          | .000         | -.228                   | .042          | .000         |
| FIELD STUDY    |                        |               | .000         |                         |               | .000         |
| FIELD STUDY(1) | .789                   | .174          | .000         | .764                    | .174          | .000         |
| FIELD STUDY(2) | .605                   | .198          | .002         | .596                    | .198          | .003         |
| FIELD STUDY(3) | .473                   | .131          | .000         | .461                    | .131          | .000         |
| FIELD STUDY(4) | .030                   | .166          | .856         | .033                    | .166          | .844         |
| FIELD STUDY(5) | .354                   | .169          | .036         | .353                    | .169          | .036         |
| FIELD STUDY(6) | -.869                  | .264          | .001         | -.871                   | .264          | .001         |
| KN THEO        | -.322                  | .029          | .000         | -.316                   | .030          | .000         |
| LEVEST P       |                        |               | .000         |                         |               | .000         |
| LEVEST P(1)    | .555                   | .145          | .000         | .563                    | .145          | .000         |
| LEVEST P(2)    | .392                   | .165          | .017         | .399                    | .165          | .016         |
| LEVEST P(3)    | .106                   | .171          | .534         | .112                    | .171          | .514         |
| LEVEST P(4)    | .218                   | .164          | .184         | .220                    | .164          | .180         |
| NUM PER        |                        |               | .286         |                         |               | .276         |
| NUM PER(1)     | -.128                  | .131          | .328         | -.131                   | .131          | .319         |
| NUM PER(2)     | -.245                  | .114          | .032         | -.248                   | .114          | .030         |
| NUM PER(3)     | -.279                  | .153          | .069         | -.279                   | .153          | .069         |
| NUM PER(4)     | -.188                  | .152          | .217         | -.190                   | .152          | .213         |
| NUM PER(5)     | -.062                  | .167          | .710         | -.060                   | .167          | .718         |
| MARK           | -.296                  | .115          | .010         | -.280                   | .115          | .015         |
| SECTOR 7       |                        |               | .000         |                         |               | .000         |
| SECTOR 7(1)    | -.378                  | .269          | .161         | -.370                   | .269          | .169         |
| SECTOR 7(2)    | -.171                  | .153          | .265         | -.174                   | .153          | .257         |
| SECTOR 7(3)    | -.573                  | .212          | .007         | -.571                   | .212          | .007         |
| SECTOR 7(4)    | .302                   | .164          | .066         | .297                    | .164          | .071         |
| SECTOR 7(5)    | -.689                  | .129          | .000         | -.688                   | .129          | .000         |

<sup>10</sup> In this regard, see references from studies such as Verdugo and Verdugo (1989), Capsada-Munsech (2015), Erdsiek (2016) or Verhaest and Verhofstadt (2016).

|                |        |      |                |        |       |      |
|----------------|--------|------|----------------|--------|-------|------|
| SECTOR 7(6)    | -1.299 | .139 | .000           | -1.301 | .139  | .000 |
| SEX(1)         | -.040  | .093 | .663           | -.036  | .093  | .695 |
| THEO PRAC KN   | -.277  | .044 | .000           | -.277  | .044  | .000 |
| TIP CONT       | -.441  | .099 | .000           | -.442  | .099  | .000 |
| TIPJORN        | .488   | .153 | .001           | .495   | .153  | .001 |
| IMR            |        |      |                | 5.446  | 3.731 | .144 |
| Constante      | -.154  | .520 | .767           | -1.921 | 1.318 | .145 |
| R2             | 0.390  |      | R2             | 0.391  |       |      |
| Classification |        |      | Classification |        |       |      |
| Overall        | 78.0%  |      | Overall        | 79.5%  |       |      |
| Normal         | 78.0%  |      | Normal         | 79.7%  |       |      |
| Overeducated   | 77.9%  |      | Overeducated   | 78.0%  |       |      |

Source: compiled by authors based on statistical analysis.

Table 5. Results of the model's explanatory variables in the second period

|                | Logit binary 2007-2011 |         |                | Heckman model 2007-2011 |               |              |
|----------------|------------------------|---------|----------------|-------------------------|---------------|--------------|
|                | B(coeficient)          | Standar | Significance   | B(coeficient)           | Standar error | Significance |
| SENIORITY      | -0.131                 | 0.038   | 0.001          | -0.127                  | 0.038         | 0.001        |
| EARNINGS       |                        |         | 0.000          |                         |               | 0.000        |
| EARNINGS(1)    | 2.838                  | 0.419   | 0.000          | 2.819                   | 0.418         | 0.000        |
| EARNINGS(2)    | 2.201                  | 0.395   | 0.000          | 2.181                   | 0.395         | 0.000        |
| EARNINGS(3)    | 2.262                  | 0,375   | 0.000          | 2.240                   | 0.375         | 0.000        |
| EARNINGS(4)    | 2.034                  | 0.371   | 0.000          | 2.016                   | 0.371         | 0.000        |
| EARNINGS(5)    | 1.415                  | 0.352   | 0.000          | 1.403                   | 0.352         | 0.000        |
| EARNINGS(6)    | 0.602                  | 0.357   | 0.092          | 0.592                   | 0.356         | 0.096        |
| EARNINGS(7)    | 0.239                  | 0.380   | 0.529          | 0.241                   | 0.379         | 0.525        |
| EST&WORK       |                        |         | 0.117          |                         |               | 0.125        |
| EST&WORK(1)    | -0.142                 | 0.116   | 0.221          | -0.144                  | 0.116         | 0.215        |
| EST&WORK(2)    | -0.244                 | 0.118   | 0.038          | -0.240                  | 0.118         | 0.042        |
| GLOB EDU       | -0.196                 | 0.035   | 0.000          | -0.188                  | 0.036         | 0.000        |
| INST KN        | 0.017                  | 0.050   | 0.727          | 0.018                   | 0.050         | 0.720        |
| INST PERS COMP | 0.098                  | 0.050   | 0.052          | 0.109                   | 0.051         | 0.034        |
| JOB SATISF     | -0.101                 | 0.045   | 0.025          | -0.099                  | 0.045         | 0.029        |
| FIELD STUDY    |                        |         | 0.000          |                         |               | 0.000        |
| FIELD STUDY(1) | 0.551                  | 0.187   | 0.003          | 0.545                   | 0.188         | 0.004        |
| FIELD STUDY(2) | 0.327                  | 0.235   | 0.164          | 0.327                   | 0.236         | 0.165        |
| FIELD STUDY(3) | 0.105                  | 0.132   | 0.425          | 0.097                   | 0.132         | 0.463        |
| FIELD STUDY(4) | -0.634                 | 0.167   | 0.000          | -0.626                  | 0.167         | 0.000        |
| FIELD STUDY(5) | 0.136                  | 0.176   | 0.439          | 0.134                   | 0.176         | 0.448        |
| FIELD STUDY(6) | -1.033                 | 0.253   | 0.000          | -1.036                  | 0.253         | 0.000        |
| KN THEO        | -0.481                 | 0.033   | 0.000          | -0.475                  | 0.034         | 0.000        |
| LEVEST P       |                        |         | 0.000          |                         |               | 0.000        |
| LEVEST P(1)    | 0.546                  | 0.144   | 0.000          | 0.541                   | 0.144         | 0.000        |
| LEVEST P(2)    | 0.797                  | 0.170   | 0.000          | 0.789                   | 0.170         | 0.000        |
| LEVEST P(3)    | 0.448                  | 0.161   | 0.005          | 0.440                   | 0.162         | 0.006        |
| LEVEST P(4)    | 0.300                  | 0.163   | 0.066          | 0.294                   | 0.163         | 0.071        |
| NUM PER        |                        |         | 0.219          |                         |               | 0.222        |
| NUM PER(1)     | -0.083                 | 0.142   | 0.561          | -0.083                  | 0.143         | 0.563        |
| NUM PER(2)     | -0.241                 | 0.123   | 0.051          | -0.240                  | 0.123         | 0.051        |
| NUM PER(3)     | -0.352                 | 0.165   | 0.033          | -0.352                  | 0.165         | 0.033        |
| NUM PER(4)     | -0.048                 | 0.167   | 0.774          | -0.054                  | 0.167         | 0.747        |
| NUM PER(5)     | -0.044                 | 0.190   | 0.819          | -0.040                  | 0.190         | 0.832        |
| MARK           | 0.114                  | 0.091   | 0.210          | 0.149                   | 0.095         | 0.116        |
| SECTOR 7       |                        |         | 0.000          |                         |               | 0.000        |
| SECTOR 7(1)    | -1.137                 | 0.356   | 0.001          | -1.124                  | 0.357         | 0.002        |
| SECTOR 7(2)    | -0.608                 | 0.184   | 0.001          | -0.594                  | 0.185         | 0.001        |
| SECTOR 7(3)    | -1.501                 | 0.312   | 0.000          | -1.506                  | 0.313         | 0.000        |
| SECTOR 7(4)    | 0.054                  | 0.180   | 0.763          | 0.059                   | 0.181         | 0.743        |
| SECTOR 7(5)    | -0.782                 | 0.157   | 0.000          | -0.766                  | 0.158         | 0.000        |
| SECTOR 7(6)    | -1.749                 | 0.164   | 0.000          | -1.735                  | 0.164         | 0.000        |
| SEX(1)         | -0.216                 | 0.102   | 0.035          | -0.213                  | 0.102         | 0.037        |
| THEO PRAC KN   | -0.193                 | 0.047   | 0.000          | -0.189                  | 0.047         | 0.000        |
| TIP CONT       | -0.348                 | 0.107   | 0.001          | -0.349                  | 0.108         | 0.001        |
| TIPJORN        | -0.035                 | 0.159   | 0.824          | -0.023                  | 0.160         | 0.888        |
| IMR            |                        |         |                | 3.001                   | 2.269         | 0.186        |
| Constante      | 0.911                  | 0.509   | 0.074          | -0.221                  | 0.994         | 0.824        |
| R2             | 0.430                  |         | R2             | 0.431                   |               |              |
| Classification |                        |         | Classification |                         |               |              |
| Overall        | 78.9%                  |         | Overall        | 78.8%                   |               |              |
| Normal         | 79.0%                  |         | Normal         | 78.8%                   |               |              |
| Overeducated   | 78.9%                  |         | Overeducated   | 79.0%                   |               |              |

Source: compiled by authors based on statistical analysis.

Table 6. Results of the model's explanatory variables in the third period

|                | Logit binary 2010-2014 |         |                | Heckman model 2010-2014 |               |              |
|----------------|------------------------|---------|----------------|-------------------------|---------------|--------------|
|                | B(coeficient)          | Standar | Significance   | B(coeficient)           | Standar error | Significance |
| SENIORITY      | -0.072                 | 0.035   | 0.038          | -0.068                  | 0.035         | 0.053        |
| EARNINGS       |                        |         | 0.000          |                         |               | 0.000        |
| EARNINGS(1)    | 2.674                  | 0.372   | 0.000          | 2.644                   | 0.372         | 0.000        |
| EARNINGS(2)    | 2.214                  | 0.362   | 0.000          | 2.190                   | 0.362         | 0.000        |
| EARNINGS(3)    | 1.634                  | 0.351   | 0.000          | 1.630                   | 0.351         | 0.000        |
| EARNINGS(4)    | 1.472                  | 0.351   | 0.000          | 1.468                   | 0.351         | 0.000        |
| EARNINGS(5)    | 0.832                  | 0.340   | 0.014          | 0.818                   | 0.340         | 0.016        |
| EARNINGS(6)    | 0.407                  | 0.346   | 0.239          | 0.408                   | 0.346         | 0.238        |
| EARNINGS(7)    | 0.036                  | 0.375   | 0.924          | 0.037                   | 0.375         | 0.921        |
| EST&WORK       |                        |         | 0.245          |                         |               | 0.238        |
| EST&WORK(1)    | -0.102                 | 0.102   | 0.319          | -0.100                  | 0.103         | 0.328        |
| EST&WORK(2)    | -0.172                 | 0.103   | 0.093          | -0.174                  | 0.103         | 0.090        |
| GLOB EDU       | -0.295                 | 0.027   | 0.000          | -0.287                  | 0.027         | 0.000        |
| INST KN        | -0.036                 | 0.044   | 0.410          | -0.034                  | 0.044         | 0.437        |
| INST PERS COMP | 0.149                  | 0.043   | 0.000          | 0.166                   | 0.043         | 0.000        |
| JOB SATISF     | -0.160                 | 0.041   | 0.000          | -0.157                  | 0.041         | 0.000        |
| FIELD STUDY    |                        |         | 0.000          |                         |               | 0.000        |
| FIELD STUDY(1) | 0.508                  | 0.204   | 0.013          | 0.484                   | 0.206         | 0.019        |
| FIELD STUDY(2) | 0.060                  | 0.229   | 0.795          | 0.035                   | 0.229         | 0.879        |
| FIELD STUDY(3) | -0.075                 | 0.124   | 0.546          | -0.083                  | 0.125         | 0.505        |
| FIELD STUDY(4) | -0.155                 | 0.140   | 0.270          | -0.163                  | 0.141         | 0.247        |
| FIELD STUDY(5) | 0.016                  | 0.175   | 0.929          | 0.002                   | 0.175         | 0.993        |
| FIELD STUDY(6) | -0.926                 | 0.211   | 0.000          | -0.938                  | 0.211         | 0.000        |
| KN THEO        | -0.390                 | 0.027   | 0.000          | -0.377                  | 0.027         | 0.000        |
| LEVEST P       |                        |         | 0.000          |                         |               | 0.000        |
| LEVEST P(1)    | 0.435                  | 0.117   | 0.000          | 0.445                   | 0.117         | 0.000        |
| LEVEST P(2)    | 0.668                  | 0.150   | 0.000          | 0.692                   | 0.150         | 0.000        |
| LEVEST P(3)    | 0.284                  | 0.128   | 0.027          | 0.298                   | 0.129         | 0.021        |
| LEVEST P(4)    | 0.162                  | 0.129   | 0.211          | 0.173                   | 0.130         | 0.183        |
| NUM PER        |                        |         | 0.175          |                         |               | 0.145        |
| NUM PER(1)     | -0.109                 | 0.123   | 0.376          | -0.108                  | 0.123         | 0.382        |
| NUM PER(2)     | -0.223                 | 0.113   | 0.048          | -0.227                  | 0.113         | 0.045        |
| NUM PER(3)     | -0.232                 | 0.144   | 0.107          | -0.233                  | 0.144         | 0.107        |
| NUM PER(4)     | -0.066                 | 0.146   | 0.652          | -0.068                  | 0.146         | 0.644        |
| NUM PER(5)     | 0.158                  | 0.174   | 0.364          | 0.177                   | 0.175         | 0.311        |
| MARK           | -0.201                 | 0.073   | 0.006          | -0.166                  | 0.075         | 0.027        |
| SECTOR 7       |                        |         | 0.000          |                         |               | 0.000        |
| SECTOR 7(1)    | -0.352                 | 0.250   | 0.159          | -0.352                  | 0.252         | 0.161        |
| SECTOR 7(2)    | -0.382                 | 0.177   | 0.031          | -0.370                  | 0.178         | 0.037        |
| SECTOR 7(3)    | -1.485                 | 0.376   | 0.000          | -1.473                  | 0.376         | 0.000        |
| SECTOR 7(4)    | 0.022                  | 0.172   | 0.899          | 0.031                   | 0.173         | 0.858        |
| SECTOR 7(5)    | -0.857                 | 0.164   | 0.000          | -0.841                  | 0.165         | 0.000        |
| SECTOR 7(6)    | -1.656                 | 0.161   | 0.000          | -1.640                  | 0.162         | 0.000        |
| SEX(1)         | 0.007                  | 0.091   | 0.938          | 0.020                   | 0.091         | 0.827        |
| THEO PRAC KN   | -0.233                 | 0.042   | 0.000          | -0.224                  | 0.042         | 0.000        |
| TIP CONT       | -0.137                 | 0.092   | 0.137          | -0.142                  | 0.092         | 0.123        |
| TIPJORN        | 0.299                  | 0.127   | 0.018          | 0.304                   | 0.127         | 0.016        |
| IMR            |                        |         |                | 5.262                   | 2.627         | 0.045        |
| Constante      | 1.217                  | 0.459   | 0.008          | -0.749                  | 1.073         | 0.485        |
| R2             | 0.479                  |         | R2             | 0.479                   |               |              |
| Classification |                        |         | Classification |                         |               |              |
| Overall        | 81.7%                  |         | Overall        | 81.7%                   |               |              |
| Normal         | 82.2%                  |         | Normal         | 82.3%                   |               |              |
| Overeducated   | 79.1%                  |         | Overeducated   | 79.0%                   |               |              |

Source: compiled by authors based on statistical analysis.

The variables related to labour market conditions exhibit the expected behaviour.

The EARNINGS variable is significant in the three periods and for all wage levels, except the two highest. The coefficients' values and signs indicate the existence of a strong negative relationship between this variable and the probability of overeducation; this relationship is accentuated for the lowest wages. However, there is a slight decrease in the variable's parameters between 2008 and 2014, probably due to wage reductions during that period: in 2008, 33 % of graduates had an income of over €24 000, while in 2014 that percentage had decreased to 23.5 %. In general terms, this result verifies those obtained by above-mentioned authors such as Duncan and Hoffman (1981), Rumberger (1987), Verdugo and Verdugo (1989), Alba-Ramirez (1993), Hartog (2000) and Rubb (2003), among others.

The variable job satisfaction (JOB\_SATISF) also exhibits a negative relationship with overeducation: greater satisfaction with the job and the theoretical knowledge used means lower probability of overeducation four years after graduation. The values of the parameters during the three periods analysed

indicate stability in the graduates' perception of their employment situation. This result verifies those obtained by Allen and van der Velden (2001), Verhaest and Omey (2009) and Diem (2015).

Therefore, classic factors such as wage level and job satisfaction play a relevant role in explaining overeducation. Indeed, a high probability of overeducation is associated with low wages and low job satisfaction. These results reveal a serious problem with job match in the labour market and partially corroborate the criticism of human capital theory for its limitations when it comes to explaining the balance between investment in education and wage returns.

The more visible credentialist variables, or the 'signals' of the academic record, and the less visible personal and instrumental skills have different values and statistical significance.

Among the 'signals', GLOB\_EDU, THEO\_PRAC\_KN and KN\_THEO are significant in the three periods with negative values, therefore they have a negative relationship with the probability of overeducation. The same goes for MARK, though this variable is not significant in the year 2011. These three variables reveal a negative relationship between the knowledge acquired and the probability of overeducation, in accordance with Spence (1973).

The variables INST\_KN and INST\_PERS\_COMP, which cover the skills learned, have a more vague and contradictory role in explaining overeducation. Both variables have positive coefficients, which indicate a higher probability of overeducation. This result indicates asymmetry with respect to the value attributed to these recruitment factors by employers and employees, characterised by the low value attributed to personal and instrumental skills by employers, compared to the high value attributed to them by graduates. However, their level of significance across the different periods is limited: the first variable is only significant in 2008, while the second is significant in 2008 and 2014.

Among the variables relating to career mobility, the influence of seniority (SENIORITY) has reduced through use of the Heckman method. It is only clearly significant in the 2007-2011 period, with a negative coefficient. This result corroborates the theoretical hypothesis, not in full, but partially; that is, in terms of career development within the firm: the greater the seniority in the post, the lower the probability of overeducation.

Also interesting is the explanation provided by the variable EST&WORK, which shows negative coefficients for EST&WORK(2), and is significant for years 2008 and 2011. This shows the importance of 'prior professional career'; that is, doing a job related to your studies during your degree reduces the probability of overeducation. This result is closer to those obtained by Sicherman and Galor (1990) and Robst (1995) in the United States than those obtained by Büchel and Mertens (2004) in Germany, and shows that career development, which provides mutual employer-employee experience and knowledge, reduces the probability of overeducation.

The role of career mobility has also lost relevance in the model because of the economic crisis and unemployment, which have altered graduates' recruitment conditions and therefore changed their career prospects.

The family socio-economic background (LEVEST\_P) is clearly significant for the three periods in cases where both parents have a primary education or one of them has a secondary education. It is also significant in cases where both have a secondary education, but only in the years 2011 and 2014. In all cases, the positive coefficients and the odds show that there is a clear negative relationship between the parents' level of studies and the probability of overeducation: the lower the parents' level of studies, the higher the probability of overeducation. Therefore, this result confirms those obtained by above-mentioned authors such as Capsada-Munsech (2015), Gaeta (2015) and Erdsiek (2016), among others.

Among the structural variables, field of study (FIELD\_STUDY) is significant in all three periods in the areas 'Humanities and Arts' and 'Health Sciences'. In these cases, the results are clear: graduates in the 'Humanities and Arts' have a higher probability of overeducation than graduates in 'Technology' (model reference), while graduates in the area 'Health Sciences' have the lowest probability of overeducation. The economic crisis, with the widespread increase in unemployment, has clearly reduced the significance of the role of field of study in recruitment; that is why this variable is less significant in 2014 than in previous periods. In terms of the variables relating to economic and firm characteristics, the results are mixed. With regard to the sectoral variables, it is worth highlighting the 'Construction' sector (SECTOR\_7(3)), the 'ICT, Finance and Business Services' sector (SECTOR\_7(5)) and 'Healthcare, Education and R&D' (SECTOR\_7(6)), whose probability of overeducation clearly reduces with regard to the reference sector: 'Public Services and Other'. In all three cases, the values of the coefficients and the odds show that these variables are relevant. Following the hypotheses of Di Pietro (2002) and Ghignoni and Vershchagina

(2014), it appears that the greater knowledge requirements of these sectors<sup>11</sup> have facilitated greater demand for graduates and their improved labour market integration, reducing the probability of overeducation. Nevertheless, the size, or firms' characteristics (NUM\_PER), have little relevance.

The economic and technological characteristics of the work environment do not play a very relevant role in explaining overeducation. The variable TIPJORN (type of working day) is significant in 2008 and 2014. In both cases, the result shows that graduates working part-time value the quality of their labour market integration more highly, since they seem to consider their job to be more efficient. In short, they have a lower probability of overeducation than those working full-time. In terms of the contract type variable (TIP\_CONT), it is significant in 2008 and 2011, but with a different influence in each case.

Lastly, sex (SEX) is only significant in 2011, with a lower probability of overeducation among women, while in most cases the characteristics of the firm are not significant. In addition, interaction effects have not been identified. The fact that there are male-dominated and female-dominated areas among the different fields of academic study probably makes it difficult for a gender bias to appear. Let us bear in mind that some of the academic areas with the lowest rates of overeducation are female-dominated areas.

The results of the statistical analysis can be used to provide a summary classification of graduates that are overeducated four years after completing their studies. Table 6 presents the main relevant factors.

Table 6. Summary of factors determining the probability of an individual being overeducated

|   |
|---|
| Labour market integration during economic crisis  |
| Low wages: below an annual income of €24 000, overeducation increases as wages decrease                             |
| Decreasing level of satisfaction, both in terms of overall job satisfaction and in the use of theoretical knowledge |
| Low level of performance in university studies  |
| Low level of theoretical-practical training and overall university training   |
| No work related to studies undertaken during degree   |
| At the start of the integration process; limited seniority in the job   |
| Family environment with a low cultural level; neither of the parents has more than a secondary education            |
| Education in the area 'Humanities and Arts'   |
| Employment in sectors with lower knowledge requirements   |

Source: Compiled by authors based on the results of the statistical analysis.

## 5. Conclusions and policy implications.

A general explanation of overeducation has been developed using the main theoretical aspects and factors on which it is based. The results obtained through the statistical analysis corroborate the role of most of the factors of overeducation relating to labour market integration. Moreover, it is worth highlighting the role of new aspects such as the sociocultural environment, which emerges as a relevant factor in explaining the probability of overeducation.

We have identified several factors that contribute to explaining overeducation; however, cyclical variables, such as unemployment or economic growth rates, were not found to be significant. Thus, the economic crisis has caused some changes related to the economic cycle, as well as the widespread increase of overeducation and job mismatch, with differing impacts according to the field of study, the role of some variables has shifted. In particular, the role of variables relating to career mobility has diminished and is less significant in 2014 than previously, while the same can be said for the field of study. Therefore, the crisis has changed the expectations of employers and employees in relation to the labour market, and the field of study as a reference in hiring.

<sup>11</sup> The 'Construction' sector appears in this group probably due to the fact that it has taken in a considerable number of graduates from technologies specifically related to this sector.

Wage level and job satisfaction play a relevant role in explaining the phenomenon, with a negative relationship between both variables and the probability of overeducation. This result shows some limitations of the human capital theory to explain the relationship between investment in education and wage returns, and its consequences on satisfaction of employees. This result is in line with several previous studies, as indicated above.

Signalling theory and the characteristics of the education received could have a positive impact on reducing overeducation, therefore validating these approaches, which originally emerged from human capital theory. Indeed, although many graduates are not rewarded for the overall investment made in their training, those who get the best results from that investment do appear to reduce their probability of overeducation. There is, therefore, differentiation between graduates based on the quality of the knowledge acquired during their studies.

Strictly academic training, which is more easily identifiable through signalling, is more highly regarded in employment integration and, as a result, high levels of training reduce the probability of overeducation. However, less visible personal skills and knowledge do not play a clear role in reducing the probability of overeducation. In this regard, graduates' labour market integration is guided by credentialism, which is related to signalling.

Despite the changes linked to the economic cycle, the hypotheses associated with career mobility theory have also been corroborated. Seniority within the firm appears to reduce the probability of overeducation in 2011, which, following authors such as Sicherman and Galor (1990) and Robst (1995), supports the importance of career development. Moreover, the role of employment activity during the final years of undergraduate studies must be highlighted. This result shows that proximity between academic activity and labour market activity, even before graduates complete their studies, can reduce the probability of overeducation. This relationship is very relevant from the point of view of the advantage of implementing work placements during the period of study.

The role of the family sociocultural environment is also noteworthy due to its positive impact on reducing overeducation. This result shows that an environment with a higher cultural level forms a better quality framework for strategic decisions in terms of facilitating graduates' choices about their labour market integration process from a short- and long-term professional perspective. In this regard, it coincides with the results of Capsada-Munsech (2015), Gaeta (2015) and Erdsiek (2016), and, from the perspective of social capital, it coincides with Verhaest and Omeij (2010). Also in this case, as in the case of career development, universities could have an impact through mentoring or guidance for labour market integration.

As regards the structural factors of overeducation or technological conditions, a clear pattern can be observed in relation to the sectors in which graduates work: sectors with greater knowledge requirements and that demand qualification credentials help reduce the risk of overeducation, as found in recent studies in this regard (Di Pietro 2002; Ghignoni and Verashchagina 2014). These are also some of the sectors with the greatest employment growth in recent decades. On the other hand, the variable corresponding to firm size shows no relevant pattern in its role in overeducation.

Also from the point of view of supply, field of study can clearly determine greater probability of overeducation, particularly for the field 'Humanities and Arts' in comparison to the rest of the fields of study, especially 'Technology'. This reveals a mismatch between supply and demand and, consequently, in the overall production system.

Finally, according to these conclusions, it is possible to act against overeducation, on the one hand, by structural measures promoting technological progress and allowing greater absorption of skilled labour by the production system. On the other hand, measures implemented in higher education itself could also have an influence, for example:

- measures to improve the match between the fields of study and the production system: innovation and technological approach;
- measures to improve guidance for students and graduates in terms of their choice of career, replacing and reinforcing sociocultural environment, and participation in work placements during their final years of study: personal and academic guidance approach.

The university could be the right place to bring together and align both approaches.



### Annex 1. Parameters of the first equation (Heckman)

|             | Period 2004-2008 |               |              | Period 2007-2011               |               |              | Period 2011-2014               |               |              |
|-------------|------------------|---------------|--------------|--------------------------------|---------------|--------------|--------------------------------|---------------|--------------|
|             | B(coeficient)    | Standar error | Significance | B(coeficient)<br>(coefficient) | Standar error | Significance | B(coeficient)<br>(coefficient) | Standar error | Significance |
| aequip      | .160             | 0.032         | .000         | .116                           | .024          | .000         | .022                           | .021          | .290         |
| apra        | -.010            | 0.032         | .762         | .063                           | .023          | .007         | .005                           | .018          | .797         |
| asolprob    | .054             | 0.032         | .091         | .120                           | .024          | .000         | .110                           | .021          | .000         |
| ateor       | .038             | 0.035         | .276         | .011                           | .026          | .684         | .068                           | .021          | .001         |
| contestu    |                  |               | .000         |                                |               | .000         |                                |               | .000         |
| contestu(1) | .922             | 0.167         | .000         | .745                           | .134          | .000         | .569                           | .133          | .000         |
| contestu(2) | .907             | 0.175         | .000         | .472                           | .136          | .001         | .184                           | .135          | .173         |
| contestu(3) | .182             | 0.168         | .279         | -.062                          | .134          | .641         | -.083                          | .137          | .543         |
| contestu(4) | .527             | 0.158         | .001         | .161                           | .119          | .174         | .137                           | .121          | .258         |
| contestu(5) | .430             | 0.213         | .044         | .492                           | .186          | .008         | .751                           | .206          | .000         |
| MARK        | .139             | 0.111         | .211         | .272                           | .069          | .000         | .134                           | .046          | .004         |
| Constant    | .814             | 0.258         | .002         | -.256                          | .201          | .204         | .371                           | .180          | .039         |

Source: compiled by authors based on statistical analysis.

### Annex 2. Principal component analysis

Rotated component matrix<sup>a</sup> 2008

|   | Component |       |        |        |
|---|-----------|-------|--------|--------|
|   | 1         | 2     | 3      | 4      |
| Satisfaction with job content               | 0.081     | 0.776 | 0.229  | -0.108 |
| Satisfaction with prospects for improvement | 0.184     | 0.737 | -0.055 | 0.160  |
| Satisfaction with wage level                | -0.015    | 0.723 | -0.053 | 0.153  |
| Overall job satisfaction                    | 0.071     | 0.807 | 0.225  | -0.070 |
| Theoretical knowledge                       | 0.004     | 0.140 | 0.838  | 0.117  |
| Practical knowledge                         | 0.302     | 0.058 | 0.735  | 0.123  |
| Language training (command of languages)    | 0.082     | 0.086 | 0.095  | 0.867  |
| Computers and new technologies              | 0.379     | 0.014 | 0.174  | 0.662  |
| Personality and social skills               | 0.857     | 0.052 | 0.075  | 0.140  |
| Ability to manage and plan                  | 0.862     | 0.121 | 0.120  | 0.155  |
| Capacity for teamwork                       | 0.855     | 0.110 | 0.116  | 0.104  |

Extraction method: principal component analysis.

a. The rotation converged in 5 iterations explaining 69.443 % of the variance of the 11 original variables.

Rotated component matrix<sup>a</sup> 2011

|  | Component |
|--|-----------|
|--|-----------|

|   | 1     | 2     | 3      | 4      |
|---|-------|-------|--------|--------|
| Satisfaction with job content               | 0.058 | 0.780 | 0.252  | -0.100 |
| Satisfaction with prospects for improvement | 0.156 | 0.732 | -0.006 | 0.170  |
| Satisfaction with wage level                | 0.008 | 0.717 | -0.071 | 0.161  |
| Overall job satisfaction                    | 0.090 | 0.824 | 0.209  | -0.066 |
| Theoretical knowledge                       | 0.029 | 0.124 | 0.846  | 0.118  |
| Practical knowledge                         | 0.289 | 0.098 | 0.730  | 0.134  |
| Language training (command of languages)    | 0.080 | 0.113 | 0.100  | 0.858  |
| Computers and new technologies              | 0.335 | 0.011 | 0.163  | 0.712  |
| Personality and social skills               | 0.860 | 0.057 | 0.081  | 0.135  |
| Ability to manage and plan                  | 0.850 | 0.110 | 0.137  | 0.156  |
| Capacity for teamwork                       | 0.844 | 0.110 | 0.118  | 0.117  |

Extraction method: principal component analysis.

a. The rotation converged in 5 iterations explaining 69.584 % of the variance of the 11 original variables.

**Rotated component matrix<sup>a</sup> 2014**

|   | Component |       |        |        |
|---|-----------|-------|--------|--------|
|   | 1         | 2     | 3      | 4      |
| Satisfaction with job content               | 0.797     | 0.032 | 0.290  | -0.045 |
| Satisfaction with prospects for improvement | 0.708     | 0.208 | -0.016 | 0.189  |
| Satisfaction with wage level                | 0.718     | 0.086 | -0.068 | 0.138  |
| Overall job satisfaction                    | 0.845     | 0.053 | 0.203  | -0.014 |
| Theoretical knowledge                       | 0.138     | 0.067 | 0.843  | 0.121  |
| Practical knowledge                         | 0.091     | 0.277 | 0.763  | 0.137  |
| Language training (command of languages)    | 0.144     | 0.081 | 0.078  | 0.857  |
| Computers and new technologies              | 0.041     | 0.271 | 0.186  | 0.749  |
| Personality and social skills               | 0.078     | 0.853 | 0.074  | 0.123  |
| Ability to manage and plan                  | 0.121     | 0.844 | 0.162  | 0.163  |
| Capacity for teamwork                       | 0.135     | 0.842 | 0.148  | 0.116  |

Extraction method: principal component analysis.

a. The rotation converged in 5 iterations explaining 70.722 % of the variance of the 11 original variables.

### Annex 3. Summary of the variables

| NAME OF VARIABLE | COMPONENTS | EXPLANATION                |
|------------------|------------|----------------------------|
| SENIORITY        |            | Years in the firm (1-4)    |
| EARNINGS         |            | Annual earnings (8 levels) |
| EARNINGS(1)      |            | <€9000                     |
| EARNINGS(2)      |            | From €9001 to 12000        |
| EARNINGS(3)      |            | From €12001 to 15000       |
| EARNINGS(4)      |            | From €15001 to 18000       |

|                |  |  |
|----------------|--|--|
| EARNINGS(5)    |  | From €18001 to 24000   |
| EARNINGS(6)    |  | From €24001 to 30000   |
| EARNINGS(7)    |  | From €30001 to 40000   |
| EARNINGS(8)    |  | >€40000 (reference)  |
| EST&WORK       |  | Activity during studies  |
| EST&WORK(1)    |  | Only studies   |
| EST&WORK(2)    |  | Work related to studies  |
| EST&WORK(3)    |  | Work unrelated to studies                                      |
| GLOB_EDU       |  | Overall university training (Likert 1-7)                       |
| FIELD_STUDY    |  | Academic area (7 areas)  |
| FIELD_STUDY(1) |  | Humanities and Arts  |
| FIELD_STUDY(2) |  | Philology  |
| FIELD_STUDY(3) |  | Social Sciences  |
| FIELD_STUDY(4) |  | Education and Communication                                    |
| FIELD_STUDY(5) |  | Sciences   |
| FIELD_STUDY(6) |  | Health Sciences  |
| FIELD_STUDY(7) |  | Technology (reference)   |
| LEVEST_P       |  | Parents' education level (5 levels)                            |
| LEVEST_P(1)    |  | Both with primary education                                    |
| LEVEST_P(2)    |  | One with secondary education                                   |
| LEVEST_P(3)    |  | Both with secondary education                                  |
| LEVEST_P(4)    |  | One with university education                                  |
| LEVEST_P(5)    |  | Both with university education (reference)                     |
| NUM_PER        |  | Firm size (number of employees)                                |
| NUM_PER(1)     |  | <10 people   |
| NUM_PER(2)     |  | From 11 to 50  |
| NUM_PER(3)     |  | From 51 to 100   |
| NUM_PER(4)     |  | From 101 to 250  |
| NUM_PER(5)     |  | From 251 to 500  |
| NUM_PER(6)     |  | >500   |
| MARK           |  | Performance (0-5)  |
| SAT_THEO       |  | Satisfaction with theoretical knowledge used                   |
| SECTOR_7       |  | Sector of activity (7 sectors)                                 |
| SECTOR_7(1)    |  | Agriculture, Extr. ind., Electr. and Gas                       |
| SECTOR_7(2)    |  | Manufacturing  |
| SECTOR_7(3)    |  | Construction   |
| SECTOR_7(4)    |  | Commerce, Repairs, Hotels and Transport                        |
| SECTOR_7(5)    |  | ICT, Finance and Business Services                             |
| SECTOR_7(6)    |  | Health and Education   |
| SECTOR_7(7)    |  | Public Services and Other (reference)                          |
| TIP_CONT       |  | Type of contract (permanent: 0 / temporary: 1)                 |
| TIPJORN        |  | Type of working day (full-time: 0 / part-time: 1)              |
| SEX            |  | Sex (man: 0 / woman: 1)  |
| Satisf1        | JOB_SATISF: Job satisfaction                     | Quantitative variable, result of principal component factoring |
| Satisf2        |  |  |
| Satisf3        |  |  |
| Satisf5        |  |  |
| Empleab1       | THEO_PRAC_KN: Theoretical/Practical knowledge    | Quantitative variable, result of principal component factoring |
| Empleab2       |  |  |
| Empleab3       | INST_KN: Instrumental knowledge                  | Quantitative variable, result of principal component factoring |
| Empleab4       |  |  |
| Empleab5       | INST_PERS_COMP: Personal and instrumental skills | Quantitative variable, result of principal component factoring |
| Empleab6       |  |  |
| Empleab7       |  |  |

Source: compiled by authors, summary of statistical variables.

## References

- Acemoglu, D. (1998). Why do new technologies complement skills? Direct technical change and wage inequality. *The Quarterly Journal of Economics*, 113(4), 1055-1089.
- Acemoglu, D. (2002). Technical change, Inequality and the labor market. *Journal of Economic Literature*, 40(1), 7-72.
- Alba-Ramirez, A. (1993). Mismatch in the Spanish Labor Market: Overeducation?. *Journal of Human Resources*, 27(2), 259-278.
- Allen, J. & van der Velden, R. (2001). Educational Mismatches versus Skill Mismatches: Effects on Wages, Job Satisfaction and On-the-Job Search. *Oxford Economic Papers*, 53(3), 434-452.
- Autor, D., Levy, F., & Murnane, R. (2003). The skill content of recent technological change: an empirical exploration. *The Quarterly Journal of Economics*, 118(4), 1279-1333.
- Arrow, K. (1973). Higher Education as a Filter. *Journal of Public Economics*, 2, 193-216.
- Baert, S., Cockx, B. & Verhaest, D. (2013). Overeducation at the start of career: Stepping stone or trap?. *Labor Economics*, 25, 123-140.
- Barone, C. & Ortiz, L. (2011). Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation. *Higher Education*, 61, 325-337.
- Berg, I. (1970). *Education and Jobs: The Great Training Robbery*. New York: Praeger Publishers.
- Becker, G. (1964). *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education*. New York: Columbia University Press.
- Büchel, F. & Mertens, A. (2004) Overeducation, undereducation, and the theory of career mobility. *Applied Economics*,  
<http://dx.doi.org/10.1080/0003684042000229532>
- Capsada-Munsech, Q. (2015). The role of social origin and field of study on graduates' overeducation: The case of Italy. *Higher Education*, 69, 779-807.
- Capsada-Munsech, Q. (2017) Overeducation: Concept, theories, and empirical evidence, *Sociology Compass*,  
<https://doi.org/10.1111/soc4.12518>.
- Caroleo, F. & Pastore, F. (2013). Overeducation at a glance: Determinants and Wage effects of the Education Mismatch, Looking at the AlmaLaurea Data. IZA Discussion Paper, no 7788. <https://www.econstor.eu/handle/10419/90000>, Accessed 20 November 2017.
- Caroleo, F. & Pastore, F. (2017). Overeducation at a Glance. Determinants and Wage Effects of the Education Mismatch Base don AlmaLaurea Data, *Social Indicators Research*, 137, 999-1032.  
<http://doi.org/10.1007/s11205-0017-1641-1>
- Chevalier, A. (2000). Graduate Overeducation in the U.K. Centre of Economics of Education. London School of Economics and Political Science.
- Clarck, B., Joubert, C. & Maurel, A. (2014). The career prospects of overeducated Americans. NBER Working Paper Series,  
<http://www.nber.org/papers/w20167.pdf>, Accessed 15 November 2017.

- Diem, A. (2015) Overeducation among graduates from universities of applied sciences: Determinants and consequences. *Journal of Economic and Financial Studies*,  
<http://dx.doi.org/10.18533/jefs.v3i02>
- Di Pietro, G. (2002) Technological change, labor markets and low-skill, low-technology traps. *Technological and Social Change*,  
[https://doi.org/10.1016/S0040-1625\(01\)00182-2](https://doi.org/10.1016/S0040-1625(01)00182-2)
- Di Stasio, V., Bol, T. & Van de Werfhorst, H.G. (2016). What makes education positional? Institutions, overeducation and the competition for jobs. *Research in Social Stratification and Mobility*, 43, 53-63.  
<http://dx.doi.org/10.1016/j.rssm.2015.09.005>.
- Doeringer, P. & Piore, M. (1971). *Internal labor markets and Manpower analysis*. Lexington D.C.: Heath and Co.
- Dolton, P. & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review*, 19, 179-198.
- Duncan, G. & Hoffman, S. (1981). The Incidence and Wage effects of Overeducation. *Economics of Education Review*, vol 1(1), 75-86.
- Erdziek, D. (2016) Overeducation of graduates: assessing the role of family background. *Journal for Labour Market Research*,  
<http://doi.org/10.1007/s12651-016-0208-y>
- Ermini, B., Papi, L. & Scaturro, F. (2017) An analysis of the Determinants of Over-Education Among Italian Ph.D Graduates, *Italian Economic Journal*, 3, 167-207.  
<http://doi.org/10.1007/s40797-017-0053-3>.
- Fleming, C. & Kler, P. (2008) I'm too clever for this job: a bivariate probit analysis on overeducation and job satisfaction in Australia, *Applied Economics*,  
<http://doi.org/10.1080/00036840600771254>
- Freeman, R. (1976). *The Overeducated American*. New York: Academic Press.
- Freeman, R. (1978). Job satisfaction as an economic variable, *The American Economic Review*, 68(2), 135-141.
- Gaeta, J.L. (2015) Was it worth it? An empirical analysis of over-education among PhD recipients in Italy. *International Journal of Social Economics*,  
<http://dx.doi.org/10.1108/IJSE-08-2013-0186>
- García Montalvo, J. (1995). *Empleo y sobrecualificación: el caso español*. Madrid: FEDEA, Documento de Trabajo 95-20.
- García Montalvo, J. & Peiró, J.M. (2009). *Análisis de la sobrecualificación y la flexibilidad laboral*. Observatorio de inserción laboral de los jóvenes 2008. Valencia: IVIE.
- Ghignoni, E. & Verashchagina, A (2014) Educational Qualifications mismatch in Europe. Is it demand or supply driven?. *Journal of Comparative Economics*,  
<http://dx.doi.org/10.1016/j.jce.2013.06.006>
- Green, F. & Zhu, Y. (2010) Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford University Papers*,  
<http://doi.org/10.1093/oup/gpq002>
- Hair, J., Anderson, R., Tatham, R. & Black, W. (1999). *Análisis Multivariante*. Madrid: Prentice Hall Iberia.
- Hartog, J. (2000). Overeducation and earnings: Where are we, where should we go?. *Economics of Education Review*, 19(2), 131-47.
- Hartog, J. & Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: overschooling? *Economics of Education Review*, 7(2), 185-194.
- Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(I), 153-162.

- Hersh, J. (1991). Education match and job match. *Review of Economic and Statistics*, 73(1), 140-144.
- Iriondo, I. & Pérez-Amaral, T. (2016) The effect of educational mismatch on wages in Europe. *Journal of Policy Modeling*,  
<http://dx.doi.org/10.1016/j.jpolmod.2015.12.008>
- Katz, L. & Murphy, K. (1992). Changes in relative wages, 1963-1987: supply and demand factors. *The Quarterly Journal of Economics*, 107(1), 35-78.
- Kucel, A. & Vilalta-Bufí, M. (2013): Job satisfaction of university graduates, *Revista de Economía Aplicada*, 61(21), 29-55.
- Leuven, E. & Oosterbeek, H. (2011). Overeducation and Mismatch in the Labor Market, IZA Discussion Paper n 5523,  
<https://econpapers.repec.org/paper/izaizadps/dp5523.htm>, Accessed 24 November 2017
- McGuinness, S. (2006). Overeducation in the Labour Market. *Journal of Economic Surveys*, 20(3), 387-418
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.
- Nieto, S. & Ramos, R. (2017) Overeducation, Skills and Wage Penalty: Evidence for Spain Using PIAA Data. *Social Indicators Research*,  
<http://doi.org/10.1007/s11205-016-1423-1>
- OECD (2006). *Innovation and Knowledge-intensive Service Activities*. Paris: OECD.
- OECD (2017). *Education at a Glance*. Paris: OECD.
- OECD (2017): *Education at a Glance*. OECD, Paris.
- Pareja-Eastaway, M. & Turmo-Garuz, J. (2013). La necesaria transformación del modelo productivo en España: el papel del territorio, *Documents D'Anàlisi Geogràfica*, 59(3), 455-479.
- Passaretta, G. & Triventi, M. (2015). Work experience during higher education and post-graduation occupational outcomes: A comparative study on four European countries, *International Journal of Comparative Sociology*, 56(3-4), 232-253.  
<http://doi:10.1177/0020715215587772>
- Pineda-Herrero, P., Agud-Morell, I. & Ciraso-Cali, A. (2016). Factors intervening in the entrance into labour market of graduates in Education Science during crisis. A study in Catalonia, *Revista de Educación*, 372, 141-168.  
<http://doi:104438/1988-592X-RE-2015-372-318>
- Reig, E. (2017). La competitividad de las regiones españolas ante la economía del conocimiento, Bilbao, Fundación BBVA.
- Robst, J. (1995). Career mobility, job match, and overeducation. *Eastern Economic Journal*, 21, 539-50
- Rubb, S. (2003) Overeducation in the labor market: a comment and re-analysis of a meta-analysis. *Economic of Education Review*,  
[http://doi.org/10.1016/S0272-7757\(02\)00077-8](http://doi.org/10.1016/S0272-7757(02)00077-8)
- Rubb, S. (2005) Overeducation, undereducation, and the theory of career mobility: a comment and a note on underemployment. *Applied Economics Letters*,  
<http://doi.org/10.1080/1350485052000337770>
- Rumberger, R.W. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources*, 22(1), 24-50.
- Sicherman, N. & Galor, O. (1990). A Theory of Career Mobility. *Journal of Political Economy*, 98(1), 169-192.
- Smith, J. & Welch, F. (1978). The overeducated American? a review article. Santa Monica, CA: RAND Corporation.  
<http://www.rand.org/pubs/papers/P6253.html>, Accessed 15 October 2017.

- Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355-374.
- Summerfield, F. & Theodossian, I. (2017). The effects of macroeconomic conditions at graduation on overeducation, *Economic Inquiry*, 55 (3), 1370-1387.  
[http://doi: 10.1111/ecin.12446](http://doi:10.1111/ecin.12446)
- Tan, E. (2014). Human Capital Theory: A holistic criticism, *Review of Educational Research*. 84(3), 411-445.  
<http://doi:10.3102/0034654314532696>.
- Thenuwara, W. & Morgan, B. (2016). Labour supply behaviour of married women in Toronto, *Journal of Economic Studies*, 43(3), 418-431.  
<http://doi.org/10.1108/JES-12-2014-0207>
- Thurow, L. (1975). *Generating inequality*. New York: Basic.
- Vatta, K., Sato, T. & Taneja, G. (2016). Indian Labour Markets and Returns to Education, *Millennial Asia*, 7(2), 107-130.  
<http://doi:10.1177/0976399616655001>
- Verdugo, R. & Verdugo, N. (1989). The impact of surplus schooling on earnings: some additional findings. *Journal of Human Resources*, XXIV(4), 629-643.
- Verhaest, D. & Omeij, E. (2006) The impact of overeducation and its measurement. *Social Indicators Research*,  
<http://doi.org/10.1007/s11205-005-4276-6>
- Verhaest, D. & Omeij, E. (2009) Objective overeducation and worker well-being: a shadow price approach. *Journal of Economic Psychology*,  
<https://doi.org/10.1016/j.joep.2008.06.003>
- Verhaest, D. & Omeij, E. (2010) The determinants of overeducation: different measures, different outcomes?. *International Journal of Manpower*,  
<http://dx.doi.org/10.1108/01437721011073337>
- Verhaest, D. & Verhofstadt, E. (2016) Overeducation and job satisfaction: the role of job demands and control. *International Journal of Manpower*,  
<http://dx.doi.org/10.1108/IJM-04-2014-0106>
- Verhaest, D. Sellami, S. & van der Velden, R. (2017). Differences in horizontal and vertical mismatches across countries and fields of study. *International Labour Review*, 156(1), 1-23.