Economia Politica (2022) 39:861–902
https://doi.org/10.1007/s40888-022-00270-y

ORIGINAL PAPER

University education, mismatched jobs: are there gender differences in the drivers of overeducation?

David Boto-García[1] · **Marta Escalonilla**[2]

Received: 16 June 2021 / Accepted: 11 May 2022 / Published online: 9 June 2022
© The Author(s) 2022

Abstract

This paper studies potential gender differences in overeducation prevalence among recent highly educated entrants to the labour market and in its underlying drivers. Using representative microdata on university graduates in Spain in the academic year 2009/2010 observed in 2014, we analyse whether males and females exhibit distinct overeducation likelihoods, distinguishing between graduates below and above 30 years of age. First, we use matching estimators based on characteristics. Next, we estimate a Heckman probit regression with interactions between the gender indicator and graduates' characteristics to study differences between males and females in the overeducation drivers. We focus on the effects of pre- and post-graduation mobility, pre- and post-graduation labour experience, on-the-job training, first job search strategies, the field of knowledge, and English language on the probability of being overeducated. We find evidence that males and females do not exhibit different overeducation risks conditional on characteristics. In general, educational credentials that act as signals of human capital are not found to be differently valued by gender.

Keywords Overeducation · University graduates · Gender differences · Human capital

JEL Classification J16 · J24 · J61

✉ David Boto-García
botodavid@uniovi.es

Marta Escalonilla
gonzalezemarta@uniovi.es

[1] Department of Economics, Faculty of Business and Economics, University of Oviedo, Campus del Cristo s/n, 33006 Oviedo, Spain

[2] Labour Economics RULE, Department of Applied Economics, Faculty of Business and Economics, University of Oviedo, Campus del Cristo s/n, 33006 Oviedo, Spain

1 Introduction

Overeducation is a widespread phenomenon that characterizes the labour market in developed countries, especially for recent tertiary graduates. It refers to the attainment of an educational level that is higher than the one the current job requires to perform it adequately (i.e. human capital surplus).[1] It is usually understood as a market phenomenon resulting from the increase in the number of university graduates and the inability of labour demand to absorb them in the short run (Chevalier & Lindley, 2009). The economic consequences of overeducation have been widely documented. Overeducation is negatively related to satisfaction (e.g. Allen & van der Velden, 2001), job productivity (e.g. Duncan & Hoffman, 1981) and workplace labour relations (e.g. Belfield, 2010), and translates into a wage penalty relative to their well-matched peers (e.g. McGuinness & Bennett, 2007). The latter is consistent with assignment theory (Sattinger, 1993) by which wages are determined by workers' human capital and job characteristics together with potential educational mismatch. Furthermore, mismatched workers experience longer unemployment spells (e.g. Ordine & Rose, 2011). Because of this, the identification of the drivers of overeducation is relevant to policy.

A growing body of research has explored the factors that distinguish overeducated from well-matched workers. Great emphasis has been placed on the effect of regional mobility, both before (e.g. Di Pietro, 2012) and after graduation (e.g. Jauhiainen, 2011), on-the-job training (e.g. Alba-Ramírez, 1993), internships (e.g. Meroni & Vera-Toscano, 2017), field of study (e.g. Barone & Ortiz, 2011) and job search strategies (e.g. Albert & Davía, 2018). However, less attention has been paid to the potential existence of gender differences in overeducation. McGuinness, Bergin, et al. (2018) show that overeducation is higher among females in most European countries. This raises important policy implications for both higher education and labour market institutions. According to the theory of differential qualification (Frank, 1978), gender differences emerge due to females' higher reluctance to migrate and their higher preference for fields of knowledge with larger overeducation prevalence (Buser et al., 2014). Nonetheless, empirical evidence on this is mixed. Whereas some scholars show females are more likely to be overeducated (Charalambidou & McIntosh, 2021; Di Pietro & Urwin, 2006; McGuinness & Bennett, 2007; Rubb, 2014), others do not find significant differences (Acosta-Ballesteros et al., 2018; Chevalier, 2003; McGoldrick & Robst, 1996). Since the college gender gap has reversed over time (Goldin et al., 2006), it seems relevant to explore whether human capital credentials are differently related to the risk of overeducation *by gender*.

[1] Overeducation is also referred to as 'educational mismatch', 'excess education' or 'overqualification'. However, it should not be confused with skill mismatch (*overskilling*), which refers to a lack (excess) of the required skills to perform the job. As shown by Green and McIntosh (2007), the correlation between the two is around 0.2. The two are conceptually different (Allen & van der Velden, 2001). Individuals in our analysis are deemed overeducated if they hold an educational level that is higher than that needed to perform the job (Duncan & Hoffman, 1981).

This paper examines gender differences in the drivers of overeducation. Our research question is whether males and females exhibit different overeducation risks and whether its drivers are different depending on gender. Unlike most studies that compare differences in the likelihood of overeducation between males and females through a dummy variable, *ceteris paribus*, we assess if personal and background characteristics are differently associated with overeducation based on gender. We study whether pre- and post-graduation mobility, pre- and post-graduation labour experience, on-the-job training, first job search strategies, field of knowledge and English language translate into distinct overeducation likelihoods for males and females. Therefore, the paper contributes to the literature by shedding new light on whether educational credentials are differently valued depending on gender.

While several studies analyse overeducation in the general population (Bar-Haim et al., 2019; Battu & Sloane, 2004; Frei & Souza-Poza, 2012; Jauhiainen, 2011), we focus on a cohort of recent university graduates. Consistent with career mobility theory (Sicherman & Galor, 1990), overeducation is more prevalent among new entrants into the labour market and decreases with age (e.g. Alba-Ramírez, 1993). As individuals get experience and on-the-job training, they move to better matched jobs. As a result, younger cohorts are more exposed to overeducation for the same educational attainment. From this perspective, our paper is close to those by Aina and Pastore (2020), Caroleo and Pastore (2018), Cattani et al. (2018), Chevalier and Lindley (2009), Di Pietro and Urwin (2006), Dolton and Silles (2008) and Ordine and Rose (2009), who study the case of young graduates.

We analyse graduates' overeducation in Spain, a country that has experienced rapid tertiary education expansion coupled with exhibiting one of the highest incidences of overeducation among young workers (McGuinness, Bergin, et al., 2018). Previous works on overeducation prevalence in Spain include those by Acosta-Ballesteros et al. (2018), Alba-Ramírez (1993), Albert and Davía (2018) and Turmo-Garuz et al. (2019). We use microdata from a representative sample of graduates observed in 2014 that graduated in the academic year 2009–2010. This graduation cohort is relevant for several reasons. Firstly, graduating during a recession translates into earnings decline in the long term, particularly for first entrants (Escalonilla et al., 2021; Oreopoulos et al., 2012). Moreover, educational mismatches between labour demand and supply have been shown to be greater during downturns (Liu et al., 2016). Therefore, young workers who graduated during the Great Recession might have a greater risk of being overeducated. Secondly, overeducation appears to have risen faster among females following the 2008 economic crisis (McGuinness, Bergin, et al., 2018), which highlights the need to study gender differences in overeducation following recessionary shocks. Thirdly, the scars and crowding-out effects from entering the labour market in a recession might be deeper for young graduates in Spain than in other European countries given the peculiarities of its labour market, which has exhibited persistent unemployment rates (Blanchard & Jimeno, 1995), low intensity of on-the-job training (Dolado et al., 2000), and a high incidence of temporary work (Dolado et al., 2002).

Despite being a widespread phenomenon, overeducation does not affect all population groups in the same way. Similar to Albert and Davía (2018) and Carroll and Tani (2015), we distinguish between graduates aged 30 years and younger at the

time of the survey. As shown in some studies (e.g. Vera-Toscano & Meroni, 2021), birth cohorts exhibit different overeducation risks because of lifetime outcomes, which can conflate the role of labour experience (McGuinness, 2006). Moreover, by doing separate analyses for the two cohorts, we reduce the heterogeneity in unobservables within each age group.

We evaluate the classic conceptualization of overeducation as excess education based on the self-assessed method (Dolton & Silles, 2008). The statistical approach (Verdugo & Verdugo, 1989) and an overeducation measure based on the earnings penalty in the spirit of Gottschalk and Hansen (2003) are used as robustness checks. Therefore, overskilling and field of study mismatch are beyond the scope of the paper. First, we use matching estimators (Abadie & Imbens, 2006) to compare whether females and males exhibit different overeducation risks conditional on their characteristics and labour market participation propensities. Next, we estimate a Heckman probit regression that deals with self-selection into the labour market considering interaction terms between the drivers of overeducation and the gender dummy.

The remainder of the paper is structured as follows. After this introductory section, a review of the related literature is presented in Sect. 2. The database, definition of the variables and summary statistics are outlined in Sect. 3. We describe the econometric modelling in Sect. 4. The estimation results are shown and discussed in Sect. 5. Finally, Sect. 6 concludes.

2 Literature review

In this section, we first review the economic theories that explain overeducation and then outline the empirical evidence for its drivers. Finally, we discuss the rationale for the existence of gender differences in the overeducation prevalence.

2.1 Conceptual framework

The labour economics literature has proposed different explanations for the overeducation phenomenon, though all are related. Within Becker's human capital framework (Becker, 1964), overeducation is the result of a temporary mismatch between firms' technology and the human capital of the labour force that vanishes in equilibrium. However, sustained overeducation could arise as a penalty for the lack of other human capital components such as training, experience or ability. Relatedly, the job mobility theory (Sicherman & Galor, 1990) postulates that workers get into overeducated job positions because at the beginning of their labour career they lack clear signals about their productivity. Workers are thus temporarily employed in jobs with lower education requirements to acquire work experience (on-the-job training) that signals their true productivity and to have greater probability of being promoted. In this viewpoint, overeducation is a short-run stepping-stone phenomenon. Once the signal becomes credible, workers are promoted to better jobs (Sicherman, 1991).

From a different perspective, the job-matching or job-shopping theory (Jovanovic, 1979) postulates that job switching takes place as a search for better quality and matched jobs. Due to the existence of frictions in the labour market, matches are imperfect and so workers need to engage in an on-the-job search in pursuit of a better job (Dolado et al., 2009). An alternative explanation for overeducation is job competition theory (Thurow, 1975), by which workers compete for jobs based on their training costs. Under this framework, overeducation could be the result of the distribution of queues within occupations and the matching between schooling and job-specific requirements (Barnichon & Zylberberg, 2019). When the market demand for qualified jobs is scarce, only a low share of qualified workers is assigned to them, the rest being allocated to jobs requiring comparatively less education. Finally, assignment theory (Sattinger, 1993) considers that workers' selection into job positions responds to a utility maximization problem that takes into account both monetary and also non-monetary job characteristics. From this viewpoint, workers could accept overeducated job positions in exchange for other job aspects.

2.2 The drivers of overeducation

There are many factors that explain why some individuals are overeducated in their jobs.[2] We block them into five groups: (i) socioeconomic factors; (ii) credentials and field of study; (iii) geographical mobility; (iv) labour experience and on-the-job training; and (v) job search strategies.

2.2.1 Socioeconomic factors

Several studies have analysed the incidence of educational mismatches based on nationality, focusing on differences among ethnic minorities. These studies typically find that overeducation is higher among non-whites (Battu & Sloane, 2004) and recently arrived migrants (Kifle et al., 2019). Regarding differences by age cohorts, Bar-Haim et al. (2019) document that tertiary education has become more necessary and less sufficient for younger cohorts to survive in a competitive labour market, which implies that younger generations are more exposed to overeducation. When it comes to family responsibilities, some studies have looked at the effect of having children. Empirical evidence on this is unconclusive: whereas some works detect a marginally significant relationship between having children and the risk of overeducation (Barone & Ortiz, 2011; Caroleo & Pastore, 2018; Jauhiainen, 2011; Rubb, 2014), some others do not detect significant effects (Acosta-Ballesteros et al., 2018; Büchel & van Ham, 2003; Devillanova, 2013; Dolton & Silles, 2008).

The family background is another important determinant of educational outcomes and has also been associated with overeducation. Research has shown that those who come from low family backgrounds are more likely to be overeducated (Barone & Ortiz, 2011; Meroni & Vera-Toscano, 2017; Turmo-Garuz et al., 2019). In

[2] A recent review of the state of the art can be found in Leuven and Oosterbeek (2011) and McGuinness et al. (2018b).

this sense, the father's education increases the speed towards a job (Ordine & Rose, 2015) and reduces the likelihood of overeducation (Ordine & Rose, 2009). This is because those who face higher schooling costs underinvest in education (Charlot & Decreuse, 2010). In the absence of valid data about family characteristics, some studies use information on whether the individual enjoyed a scholarship for tertiary education, finding that it is positively related to overeducation (Verhaest & Omey, 2010).[3] In this vein, evidence by Denning et al. (2018) shows that grants for college students increase the probability of degree completion and later earnings.

2.2.2 Credentials and field of study

Since firms cannot directly observe workers' abilities, all type of credentials that help signal a worker's human capital play a role in lowering the overeducation prevalence. Graduating with distinction (high final marks) has been shown to increase weekly earnings after college graduation (Khoo & Ost, 2018) and to reduce the likelihood of overeducation (Albert & Davía, 2018; Devillanova, 2013; Di Pietro & Urwin, 2006; Turmo-Garuz et al., 2019; Verhaest & Omey, 2010). In this sense, Agopsowicz et al. (2020) document that graduates with low grade point averages (GPA) are less likely to hold a college job. The same applies to years of schooling, those graduating later than expected being more exposed to overeducation (Caroleo & Pastore, 2018; Ordine & Rose, 2009). Another relevant credential is knowledge of a second language, especially English. Albert and Davía (2018) find that a good level of English protects Spanish graduates from overeducation.

Universities and vocational programmes also play a role in students' acquisition of human capital. Ordine and Rose (2009, 2011) point to universities instructional quality as a major determinant of overeducation. Their public or private character might be important due to their different financial resources. Additionally, Verhaest et al. (2018) point out that programmes that combine theoretical knowledge with specific workplace learning provide new labour entrants with qualifications that are immediately usable, thereby reducing educational mismatches.

Another issue that has attracted attention is whether overeducation is more concentrated in some academic fields. Barone and Ortiz (2011), Dolton and Silles (2008), Meroni and Vera-Toscano (2017), and Turmo-Garuz et al. (2019) find that graduates from humanistic areas are at a higher risk of being overeducated. In contrast, graduating from scientific, health and technical disciplines protects workers from overeducation (Acosta-Ballesteros et al., 2018; Aina & Pastore, 2020). This has been explained by the fact that technical fields of knowledge may put more emphasis in their curricula on the acquisition of specific skills (Reimer et al., 2008). Nevertheless, skills obsolescence also matters, since fast-changing occupations like STEM fields continually demand new abilities (Deming & Noray, 2020).

[3] Since only students from low-income families receive such a grant, this variable is a proxy for family economic constraints.

2.2.3 Labour experience and on-the-job training

According to the job mobility theory, on-the-job training and internships help workers both to signal their productivity levels and to invest in occupation-specific human capital (Barron et al., 1989), so that overeducation prevalence should smoothly reduce with labour experience. However, empirical evidence on this is inconclusive. On the one hand, some scholars present evidence that the lack of experience of new labour market entrants makes them more exposed to overeducation (e.g. Sloane et al., 1996) and that experience is negatively related to the risk of overeducation (Kiker et al., 1997; Robst, 2008). On the other hand, other studies do not find significant differences (Dolton & Silles, 2008; Verhaest & Omey, 2010).

2.2.4 Geographical mobility

There is inconclusive evidence on the career benefits of transnational educational mobility (i.e. stays abroad during the completion of studies). A well-known example of an exchange programme is Erasmus. On the one hand, internationally mobile students exhibit better transitions to employment (e.g. Di Pietro, 2012) and achieve higher incomes after graduation (e.g. Kratz & Netz, 2018). This is due to the acquisition of soft skills related to cultural understanding, open-mindedness and sociability during international experiences (Crossman & Clarke, 2010). These social skills are highly valued in the labour market (e.g. Deming, 2017). What is more, those who engaged in Erasmus are more likely to move to other regions to find a better job match (Krabel & Flöther, 2014) and to work abroad later in life (Oosterbeek & Webbink, 2011; Parey & Waldinger, 2010). In this sense, Albert and Davía (2018) find that having studied abroad is negatively related to overeducation. However, other studies find that individuals who stayed abroad need more time to find a first job after graduation and are more likely to be overeducated (e.g. Wiers-Jenssen & Try, 2005), possibly due to problems with human capital transferability to local labour markets.

Since there are regional differences in the incidence of overeducation (Lenton, 2012), one way to overcome labour demand and supply mismatches is through mobility, either domestically or abroad. Jauhiainen (2011) shows that long-distance migration reduces the probability of overeducation. However, evidence on this is not robust. Devillanova (2013) finds that short distance mobility is negatively correlated with overeducation but emphasizes the need to control for workers' characteristics. In this respect, a large body of research finds that immigrants and minority ethnic groups are more likely to be overeducated (Beckhusen et al., 2013). As such, mobility increases the likelihood of migrant graduates being matched but at the cost of higher overeducation exposure in the short run.

Another relevant factor is past job turnover. In line with the job-matching theory (Jovanovic, 1979), overeducated workers are those with shorter average job durations and a higher turnover (Alba-Ramírez, 1993). Rubb (2013) finds that it might be optimal for overeducated workers to engage in firm-switching behaviour for occupational mobility. Similar results are reported by Romanov et al. (2017), who

show that the number of employers in the past three years is negatively related to the likelihood of overeducation.

2.2.5 Job search strategies

Although some studies show that overeducation is transient and is the first step to climb the occupational ladder (Frei & Souza-Poza, 2012), others find that being overeducated in the first job positively influences being mismatched in subsequent occupations (Acosta-Ballesteros et al., 2018; Baert et al., 2013; Meroni & Vera-Toscano, 2017). Therefore, first employment characteristics can exert an impact on the long-run prevalence of overeducation. In light of this evidence, another stream of research has explored the role of universities and third-party job placement institutions as intermediaries that connect recent graduates with firms. Carroll and Tani (2015) and McGuinness et al. (2016) report that job placement assistance by education institutions reduces overeducation incidence in graduates' first jobs. Concerning the use of the internet for getting the first job, these authors show it is positively related to overeducation. Furthermore, old work networks have been shown to improve labour market opportunities for newly entrants (Simon & Warner, 1992). Albert and Davía (2018) report that those who resort to temporary work agencies have a higher risk of underemployment and overeducation. The same finding holds for mass media and the internet. However, informal job search methods such as contacting employers directly or getting a job through friends and relatives are not related to skills' underutilization.

2.3 Gender differences

The literature has explored gender differences in overeducation with mixed findings. Whereas most studies show that females are more likely to be overeducated (Barone & Ortiz, 2011; Charalambidou & McIntosh, 2021; Di Pietro & Urwin, 2006; Jauhiainen, 2011; Meroni & Vera-Toscano, 2017; Ordine & Rose, 2009; Rubb, 2014), others do not find significant differences (Acosta-Ballesteros et al., 2018; Aina & Pastore, 2020; Caroleo & Pastore, 2018; Chevalier, 2003; Devillanova, 2013; Dolton & Silles, 2008; Kiker et al., 1997). One argued reason for a gender gap in overeducation is the theory of differential overqualification (Frank, 1978). According to this theory, females are more exposed to overeducation because they are less mobile due to family responsibilities. However, McGoldrick and Robst (1996) do not find evidence supporting this.

Another source of gender differences could stem from differences in career-prospects. Robst (2007) argues that males and females have different reasons for accepting a mismatched job position, and this directly relates to career expectations and preferences. Redmond and McGuinness (2019) document that females prefer jobs that are closer to home and offers job security, whereas males are more motivated by financial gains and promotion. A large body of literature has documented gender differences in risk-taking behaviour (e.g. Jetter & Watter, 2020), competitiveness (e.g. Gneezy et al., 2003) and ambition (e.g. Chevalier, 2007) in favour of males. This has

been thought to explain the gender wage gap (e.g. Le et al., 2011), deterrence factors in relation to the job entry decision for females (e.g. Flory et al., 2015) and occupational choice (e.g. Kleinjans, 2009).

Furthermore, occupational sorting and gender segregation might contribute to explaining potential gender differences in overeducation. Females continue to be underrepresented in STEM fields and prestigious academic tracks (Buser et al., 2014). Recent evidence by Barone and Assirelli (2020) points out that early curricular track choices affected by peers' preferences strongly impact the subsequent choice of field of study. Females' engagement in STEM is affected by the share of females that graduate from STEM fields (Griffith, 2010). Moreover, females are more likely to leave the STEM field for majors that are less competitive (Astorne-Figari & Speer, 2019) and to prioritize the work-life balance (Jiang, 2021).

Finally, gender differences in overeducation could arise due to taste-based or statistical discrimination by employers (Guryan & Charles, 2013). Taste-based discrimination refers to the employer *consciously* prioritizing offering a matched job to a male due to biased preferences, even though they have the same human capital. Statistical discrimination implies that statistical information about the group to which the worker belongs is used to infer his/her unobserved productivity. For instance, Deming et al. (2016) show that employers strongly prefer applicants with degrees from public institutions because they negatively value credentials from for-profit institutions. Oreopoulos (2011) reports that job applicants with foreign names and backgrounds are much less likely to receive a call back for a job interview. Additionally, gender differences could be explained by implicit discrimination (e.g. Bertrand et al., 2005), by which employers hold unconscious negative attitudes towards the capacity of females, especially in certain job positions. This relates to the existence of stereotypes, which shape beliefs about the capacities of others in different domains (Bordalo et al., 2019).

3 Data

3.1 Database

Our analysis is based on the 2014 Labour Insertion of University Graduates Survey conducted by the Spanish National Statistics Institute. This dataset is representative of all graduates in the Spanish university system during the 2009/2010 academic year. The data were collected between September 2014 and February 2015, leaving a minimum period of four years since the completion of university studies. The sample has a cross-sectional structure and consists of 30,379 observations.

As mentioned, this graduation cohort is of particular relevance because university graduates that enter the labour market in the middle of the Great Recession are the most exposed to overeducation in their jobs. Overeducation has been argued to arise due to mismatches between the labour supply and demand for educated workers (O'Leary & Sloane, 2016). During economic downturns, highly educated workers compete for a lower number of qualified job positions through underbidding the wages they demand (Leuven & Oosterbeek, 2011), thereby making new entrants

during recessions more exposed to suffer overeducation in their jobs.[4] In this regard, several studies have shown that the poor performance of the youth labour market in Spain in terms of overeducation and high unemployment is partially due to highly educated workers crowding-out lower educated ones because of the excess supply of college graduates (Dolado et al., 2000, 2009).

We follow Albert and Davía (2018) and Carroll and Tani (2015) and split the sample between graduates aged 'less than' and 'equal or greater than' 30 years at the time of the survey.[5] They are denoted 'less than' and 'over 30' throughout the paper, those exactly 30 years old being included in the latter group. Consistent with the theory of career mobility (Sicherman & Galor, 1990), the literature typically finds that older workers are less likely to be overeducated (Rubb, 2003). Beyond this, recent studies like that by Bar-Haim et al. (2019) show that younger cohorts face greater labour market competition, plausibly due to a gradual upgrading of schooling requirements. Accordingly, we conduct separate analyses to inspect whether overeducation prevalence differs across birth cohorts.[6]

3.2 Measuring overeducation

There is debate in the literature about the appropriate way to measure overeducation (Verhaest & Omey, 2006). Given the worker's educational level, the main problem is how to define the required education to perform a certain job. Three alternatives have been proposed.

1. Job analysis (JA): This procedure consists of developing 'dictionaries' of the different occupations by which experts value the education required for each job (Rumberger, 1987). Studies that rely on the JA approach include those by Acosta-Ballesteros et al. (2018), Baert et al. (2013) and Sellami et al. (2017). However, apart from being expensive and demanding, this approach is likely to become obsolete if the dictionary is not frequently updated (Hartog, 2000; McGuinness, 2006), especially in those occupations that are related to new technologies. Moreover, the same job occupation might have different educational requirements across industries and firms.
2. Individual self-assessment (ISA): Workers are directly asked to report the necessary education level 'to do' their job. Obviously, this method has the drawback of being self-reported, which implies that respondents might consider their own expectations and disappointments (Cattani et al., 2018; Verhaest & Omey, 2010).

[4] As opposed to economic downturns, educated workers should theoretically face lower difficulties finding matched jobs during economic expansions because of labour demand shifts to the right for a relatively constant supply of new educated workers. Nevertheless, Spain has exhibited large persistence in overeducation even during expansionary periods, in part due to labour market rigidities and the expansion of low-skill service employment (e.g. Consoli & Sánchez-Barrioluengo, 2019).

[5] For confidentiality reasons, the dataset does not provide information on the exact age, only if the graduate is below or over 30, so we use the age cut-off provided in the survey.

[6] Since the average age at which females have their first child is 31 years old in Spain (INE, 2019), this distinction is also relevant for potential family responsibilities.

As discussed in Hartog (2000), individuals might exhibit a tendency to overstate the job requirements to upgrade the status of their position. Moreover, this method could be subject to qualification inflation in firms' hiring strategies in the case of newly hired workers (Barone & Ortiz, 2011). Nonetheless, it is the most used (Alba-Ramírez, 1993; Chevalier, 2003; Dolton & Silles, 2008; Ordine & Rose, 2009).[7]

3. The statistical approach (SA): This method defines overeducation by comparing the worker's educational level with the mean (or mode) value within a given occupation. Some examples include Charalambidou and McIntosh (2021), Kiker et al. (1997) or Verdugo and Verdugo (1989). Although it is easy to implement, the SA merely reflects the credentials of all workers in a job position. Therefore, it can gather the education level 'to get' the job rather than 'to do' it (McGuinness, Pouliakas, et al., 2018). Other flaws include the aggregation of broad occupational groups and the fact that occupation averages might reflect historical entry requirements. Furthermore, because overeducation within occupations is the result of the interaction between demand and supply forces, this method does not reflect job requirements only (Leuven & Oosterbeek, 2011).[8]

The three approaches have advantages and weaknesses. Verhaest and Omey (2010) examine how the effect of the determinants of overeducation is sensitive to the measure used. They show that objective and subjective assessments can lead to different results, but the JA does not result in more reliable outcomes than the ISA. A similar discussion is provided in Capsada-Munsech (2019) and McGuinness (2006). The use of one or another typically depends on the information available. Nevertheless, there is wider acceptance of the ISA, as illustrated by Capsada-Munsech (2019). The reason is that it measures 'genuine' overeducation (Chevalier, 2003) rather than 'apparent' one because it exploits individual perceptions about the educational requirements to perform the job.

3.3 The self-assessment approach (ISA)

In the survey, respondents are asked the following:

Which do you think it is the most suitable educational level to perform your current job?

This question mimics the one implemented in Alba-Ramírez (1993) and Dolton and Silles (2008) for assessing self-reported overeducation in Spain and the UK, respectively. It is also very close to the one implemented in the REFLEX survey (Barone & Ortiz, 2011; Capsada-Munsech, 2019). This question captures

[7] McGuinness et al., (2018a, 2018b) shows that the subjective method is used in 149 estimates of overeducation, the statistical approach in 103 and the job analysis in 44.

[8] Another empirical problem is how to define the threshold over which an individual is overeducated. Barone and Ortiz (2011) consider an educational level above that of 80% of the people in the same occupation, whereas in Robst (2008) and Jauhiainen (2011) overeducation exists if the associated number of years of schooling is above one standard deviation the mean.

educational requirements 'to do' the job. This is different from the job recruitment standards used in Duncan and Hoffman (1981) and collected in the PIAAC survey, which could relate to screening issues. Answers to this question can be any of the following: (a) doctorate, (b) university degree, (c) higher national diploma, (d) certificate of higher education, or (e) compulsory education. This variable is denoted by *EDUC_REQ*.

To create an indicator of educational mismatch, we define a variable that compares the self-assessed educational level to perform the job adequately (*EDUC_REQ*) with actual educational level (*EDUC_LEVEL*) as follows:

$$EDUC_M = \begin{cases} 1 & if\ EDUC_REQ < EDUC_LEVEL \\ 0 & if\ EDUC_REQ = EDUC_LEVEL \\ -1 & EDUC_REQ > EDUC_LEVEL \end{cases}$$

Positive educational mismatch (overeducation) emerges when the required level of education is *lower* than that attained. Conversely, there is negative educational mismatch (undereducation) when the required education is *higher* than that attained. The latter case ($EDUC_M = -1$) is not considered here since the share of individuals in this situation is very low (1.4%). These graduates are excluded, and so our final sample comprises 26,807 observations. Our dependent variable is therefore a dummy (*OVEREDUC*) that takes value 1 if $EDUC_M > 0$ and 0 otherwise. Nonetheless, in subsection 5.4 we use alternative dependent variables for robustness: (i) the *degree* of overeducation through an ordered indicator (labelled *OVEREDUC_DEGREE,* see Supplementary Material, Section A), (ii) the statistical approach (labelled *OVEREDUC_SA,* see Supplementary Material, Section F) and (iii) an alternative overeducation indicator based on wages differences within occupations (labelled *OVEREDUC_EARNINGS*, see Supplementary Material Section G) in the spirit of Gottschalk and Hansen (2003).

Our definition of overeducation has an important advantage over other self-assessed methods used in the related literature. For instance, unlike the European Community Household Panel that asks workers to directly indicate if they feel they are qualified to do a more demanding job, overeducation is defined here as a comparison between the reported educational level required to perform the job and the one the respondent holds. That is, as opposed to directly asking whether they feel they are overeducated in their job, the elicitation is more indirect and minimizes the risk of misinterpretation as discussed in Capsada-Munsech (2019).

3.4 Explanatory variables

Here we describe the explanatory variables to be used in our analysis.

- Sociodemographic characteristics (*SocDem*): To explore the potential gender gap in overeducation, we initially include a dummy variable for females (*FEMALE*). We also consider a dummy for whether the graduate has Spanish nationality (*SPANISH*).

- University studies characteristics (*Univ*): We include a dummy for whether the university from which the individual graduated is public (*PUB_UNI*). We also take into account the field of knowledge to which they belong: sciences (*SCIENCES*), arts and humanities (*ARTS*), engineering and architecture (*ENGINEERING*), and health sciences (*HEALTH*). Social sciences (*SOCIALS*) is the base category. Since the data include university graduates from academic degrees with different lengths, we also control for the years of tertiary schooling through the variable *STU_LENGTH*.[9] Additionally, to proxy graduates' marks, we define two dummy variables relative to the attainment of grants during their university stage: (i) a general scholarship for university students (*GRANT_STU*); and (ii) a scholarship of excellence for those with the highest marks (*GRANT_EXC*). Since the latter is only awarded to those who hold outstanding grades, this variable partially captures ability in the form of cognitive intelligence and other soft skills like diligence.[10]
- Additional studies (*AddStu*): The question about the required level of education to perform the job (*EDUC_REQ*) does not distinguish between types of university certificate (graduate, Master's, etc.). It might happen that those with a Master's degree declare themselves overeducated because their education level is higher. To control for this, we include a dummy for whether the individual has completed a Master's degree (*MASTER)*. In the same vein, we consider a dummy for whether the graduate has other vocational training courses (*VOC_TRAINING*) apart from university studies, and whether they are currently studying (*CONT_STUDY*).
- English language (*English*): A good level of English is expected to protect graduates against overeducation. To proxy it, we define two dummy variables for whether the individual has (i) a certificate in English from Cambridge (*CAMB_CERTIFICATE*) or (ii) a certificate in English from the Official Language School (*OLS_CERTIFICATE*).[11]
- Pre-graduation work experience (*PregradExp*): Other variables related to work experience prior to graduating are also included. In particular, we consider dummies for whether the respondent (i) worked while she was studying (*WORK_STU*); (ii) performed work practice before graduation (i.e., internships as part of the study program, denoted by *CURRIC_PRACT*); and (iii) performed work practice by the end of her university studies (i.e., non-curricular practices, labelled *NONCURRIC_PRACT*).

[9] The sample includes graduates from a short-cycle Bachelor's degree (3 years in length), Diploma of Higher Education (at least 3 years), a pre-Bologna long-cycle university degree (at least 5 years), a Bologna university degree (at least 4 years), an engineer's degree (at least 4 years), or a degree in medicine or architecture (at least 6 years).

[10] To get a general scholarship in Spain, university students must comply with two requirements: (i) household income needs to be below a certain threshold; and (ii) students must pass a minimum number of university credits per year with a minimum average mark to continue receiving the funding. Hence, *GRANT_STU* also controls for family background and household income.

[11] The Cambridge Assessment English is a well-recognized set of exams that prove English language ability.

- Pre-graduation mobility (*PregradMob*): To control for the completion of stays in other universities, we define two dummy variables: *NAT_STAY* for staying at another Spanish university, and *ERASMUS* if the graduate participated in an Erasmus mobility programme.
- Post-graduation work experience (*PostgradExp*): We include a dummy for whether the graduate has been working for at least two years since graduation (*EXP_2YEARS*). As a proxy for job turnover, we also add a variable for the number of employers she has had (*N_EMPLOYERS*). Implicitly, the latter variable gathers whether the graduate remains in her first job.
- Post-graduation mobility (*PostgradMob*): To control for graduates' mobility *after* graduation, we define *LAB_MOB* if the individual moved to another Spanish region for labour reasons, and *ABROAD_RESIDENCE* if the graduate has lived abroad.
- Job search strategies (*JSStrat*): We consider information on how graduates looked for their first job. Specifically, we add four binary variables for whether the respondent resorted to: (i) newspaper or internet advertisements (*JS_ADV*), (ii) temporary employment agencies (*JS_TEA*), (iii) university platforms (*JS_UNI*) or (iv) work or family networks (*JS_NETWORKS*). The base category involves other methods like self-employment or public competition.
- Job rejection (*JobRej*): To control for heterogeneity in career prospects, we define a binary variable for whether the respondent has rejected a job because they considered it unsuitable (*REJECT_JOB*).
- Regional fixed effects (*RegFE*): Variations in overeducation prevalence might be affected by regional factors like the size of the regional labour market (Büchel & van Ham, 2003; McGuinness, 2006). Indeed, McGuinness, Bergin, et al. (2018) point out that gender differences in overeducation across Europe relate to country-specific dimensions. To capture this, we include a full set of regional fixed effects (autonomous community).

3.5 Summary statistics

Summary statistics of the variables are presented in Table 1, for the whole sample and separately for males and females. The last column reports a t test for the mean equality. In general, graduates' sample characteristics differ by gender. About 77 per cent of females are working compared to 81 per cent of males. Conditional on working, 24 per cent of male and 27 per cent of female graduates are over-educated in their jobs.[12] This difference is statistically significant, suggesting that females are more overeducated in the sample. As for age, while only 49 per cent of males are below 30, in the case of females this percentage is 64 per cent.[13] Likewise, we

[12] These figures are in line with previous studies, either for Spain or for other European countries. Table A3 in Supplementary Material provides an overview of incidence rates in previous literature. The variability across studies is partially due to the country, period and sample characteristics considered.

[13] As discussed in Goldin et al. (2006), the traditional college gender gap has reversed in the last decades because of females' greater incentives to invest in human capital and not to depend on their partners, and potentially lower opportunity costs of attendance to college education, among other factors.

Table 1 Descriptive statistics

Variable		All	Males	Females	t test
Dependent variables					
WORK	Dummy, 1 = the individual works	0.76	0.81	0.77	6.75***
OVEREDUC	Dummy, 1 = overeducated	0.26	0.24	0.27	− 4.59***
Explanatory variables					
Sociodemographic characteristics (SocDem)					
FEMALE	Dummy, 1 = female	0.60			
SPANISH	Dummy, 1 = has Spanish nationality	0.95	0.95	0.95	1.50
AGE_LESS30	Dummy, 1 = less than 30 years old	0.58	0.49	0.64	− 24.68***
University studies characteristics (Univ)					
PUB_UNI	Dummy, 1 = public University	0.85	0.85	0.87	− 4.48***
SCIENCES	Dummy, 1 = sciences field of knowledge	0.09	0.08	0.10	− 5.08***
ARTS	Dummy, 1 = arts/humanities field of knowledge	0.10	0.08	0.11	− 7.04***
ENGINEERING	Dummy, 1 = sciences field of knowledge	0.22	0.40	0.11	59.03***
HEALTH	Dummy, 1 = health sciences field of knowledge	0.13	0.07	0.17	− 23.79***
SOCIALS	Dummy, 1 = social sciences field of knowledge (ref. cat)	0.44	0.36	0.50	− 23.53***
GRANT_STU	Dummy, 1 = benefited from general scholarship	0.35	0.31	0.39	− 13.27***
GRANT_EXC	Dummy, 1 = benefited from scholarship of excellence	0.06	0.07	0.06	4.27***
STU_LENGTH	Length of university studies	4.12	4.18	4.09	6.89***
Additional studies (AddStu)					
MASTER	Dummy, 1 = has a Master's degree	0.33	0.34	0.34	1.62
VOC_TRAINING	Dummy, 1 = has other vocational training courses	0.13	0.12	0.14	− 3.20***
CONT_STUDY	Dummy, 1 = continues studying (another degree, a vocational training course, etc.)	0.22	0.22	0.22	0.403
English language (English)					
OLS_CERTIFICATE	Dummy, 1 = has a certificate in English from the Official Language School	0.20	0.17	0.22	− 9.81***
CAMBR_CERTIFICATE	Dummy, 1 = has a certificate in English from Cambridge	0.29	0.29	0.29	0.23

Table 1 (continued)

Variable		All	Males	Females	t test
Pre-graduation work experience (PregradExp)					
WORK_STU	Dummy, 1 = worked while studying at university	0.62	0.67	0.60	11.52***
CURRIC_PRACT	Dummy, 1 = undertook curricular practises	0.64	0.54	0.71	− 29.07***
NONCURRIC_PRACT	Dummy, 1 = undertook non-curricular practises	0.31	0.31	0.32	− 2.22**
Pre-graduation mobility (PregradMob)					
NAT_STAY	Dummy, 1 = stayed at another Spanish university for a year different from the one from which she graduated	0.09	0.11	0.08	7.61***
ERASMUS	Dummy, 1 = participated in an Erasmus program during university studies	0.08	0.09	0.08	2.22**
Post-graduate work experience (PostgradExp)					
EXP_2YEARS	Dummy, 1 = has at least 2 years of labour experience	0.69	0.72	0.68	8.17***
N_EMPLOYERS	Number of employers	2.6	2.37	2.76	− 15.66***
Post-graduate mobility (PostgradMob)					
LAB_MOB	Dummy, 1 = moved to another Spanish region after graduation for labour purposes	0.16	0.18	0.15	5.08***
ABROAD_RESIDENCE	Dummy, 1 = lived abroad after graduation	0.16	0.17	0.16	3.21***
Job search strategies (JSStrat)					
JS_ADV	Dummy, 1 = looked for first job through advertisements	0.32	0.32	0.32	− 0.36
JS_TEA	Dummy, 1 = looked for first job through temporary employment agencies	0.10	0.09	0.11	− 5.56***
JS_UNI	Dummy, 1 = looked for first job through the university	0.21	0.19	0.23	− 9.14***
JS_NETWORKS	Dummy, 1 = looked for first job through work/family networks	0.37	0.37	0.37	0.16
Job Rejection (JobRej)					
REJECT_JOB	Dummy, 1 = rejected a job because of being unsuitable	0.31	0.36	0.28	13.74***
Observations				26,807	

Table 2 Summary statistics of dependent variables by gender and age cohort

	Less 30		Over 30	
	Females	Males	Females	Males
WORK	0.772	0.791	0.771	0.821
OVEREDUC	0.256	0.222	0.297	0.262
n	10,382	5285	5732	5408

observe significant differences in the gender composition by field of study. Males are mainly concentrated in engineering (40%) and social sciences (36%). In contrast, there is a higher percentage of females in social sciences (50%) and health (17%). About 35 per cent of graduates have benefited from a general scholarship during their studies. However, only 6 per cent have achieved a scholarship of excellence.

The average length of the university studies is around 4 years. Around 33 per cent have completed a Master's degree and 13 per cent have additional vocational training studies. Interestingly, the share of females who have a certificate in English from the Official Language School is significantly greater than that of males (22% vs 17%). In contrast, the share of graduates with a certificate in English from Cambridge is the same for box sexes (29%). Pre-graduation educational mobility is quite low: 11 per cent of males and 8 per cent of females in the case of stays at another Spanish university and 9 per cent of males and 8 per cent of females took part in an Erasmus program. By contrast, post-graduation labour mobility is somewhat larger (16%). Regarding work experience, about 62 per cent of graduates combined studying and working, with 69 per cent having more than two years of labour experience. The average number of past employers is almost three. Remarkably, 71 per cent of females undertook curricular practices compared to 54 per cent of males. The most used job search strategies for first employment are work/family networks (37%) and job advertisements (32%). Finally, 36 per cent of males and 28% of females rejected a job because they considered it unsuitable.

Table 2 distinguishes the sample sizes, proportion of employed graduates and overeducation prevalence for males and females separately by age cohort. Employability rates are higher among males, with almost no differences between those aged less than 30 and over 30 in the case of females. Overeducation is four percentage points higher among graduates over 30 for both sexes (29% vs 25% for females and 26% vs 22% for males). Since all the sample graduated in 2010, this result could reflect that those over 30 graduated later than expected, which is associated with greater overeducation exposure (Caroleo & Pastore, 2018; Ordine & Rose, 2009).

4 Empirical strategy

4.1 Matching estimators

To evaluate the existence of gender differences in overeducation, we first use matching estimators based on covariates (Abadie & Imbens, 2006). Males and females

are likely to exhibit distinct labour market participation propensities because of different reasons, which might induce selection effects into the analysis of their different overeducation risks. To deal with this, we match employed and non-employed graduates by the time of the survey based on characteristics using propensity score matching (PSM). In doing so, we impose the common support condition. Next, we implement Inverse Probability Weighting Regression Adjustment (IPWRA), which consists of running separate regressions for males and females by reweighting on their labour market participation propensity scores.[14] In this way, we compute the average gender difference in conditional-on-covariates overeducation prevalence while considering males and females' potential differences in employability based on their endowment of observed characteristics. This is done both for the pooled sample and separately for the less than 30 and over 30 cohorts.

4.2 Heckman probit

Having analysed whether males and females have a different average prevalence of being overeducated, the subsequent step is to examine the factors that could drive it. Therefore, we move to regression analysis that studies the differences in the drivers of overeducation by gender.

Overeducation is only observed for the subsample that works. Individuals are likely to self-select into the labour market based on observed and unobserved characteristics. If the unobservable characteristics that drive participation correlate with the error term in the overeducation equation, then the estimates from the probit suffer from selectivity bias. Accordingly, we propose a probit model with sample selection (Heckman probit), as done by Büchel and van Ham (2003), Jauhiainen (2011), Ordine and Rose (2009), and Rubb (2014). This model comprises two equations that are jointly estimated: a binary probit for the probability of working (selection equation) and a probit model for *OVEREDUC* for those who participate in the labour market (outcome equation).

Because our purpose is to examine whether human capital credentials are differently valued by gender, in both equations we consider interaction terms between the explanatory variables and the dummy for being a female. In this way, the model allows for gender heterogeneity in the determinants of the labour participation decision and the overeducation prevalence.[15]

[14] IPWRA estimators are double-robust to misspecification since they remain consistent if either the propensity scores or the conditional mean of the outcome is misspecified. See Sloczynski and Wooldridge (2018) for further details.

[15] Alternatively, we could run separate regressions for males and females. However, that approach would be inappropriate because in such a case the coefficient estimates would be measured on different scales and would not be directly comparable. In any case, this is performed as a robustness check (see Sect. 5.4).

The model structure of the Heckman probit is:

1. *Selection equation*:

$$WORK_i^* = \gamma_1 Z_i + \gamma_2 Z_i xFEMALE_i + v_i \tag{1}$$

where $WORK_i^*$ is the latent utility from the market perspective of hiring individual i, Z_i is a vector of exogenous explanatory variables, γ_1 and γ_2 are vectors of coefficients to be estimated and v_i is a random error term.

2. *Outcome equation*:

$$OVEREDUC_i^* = \beta_1 X_i + \beta_2 X_i xFEMALE_i + u_i \tag{2}$$

where $OVEREDUC_i^*$ is a latent variable that measures the gap between an individual's level of education and the required level of education for the job, X_i is a vector of exogenous explanatory variables, β_1 and β_2 are vectors of coefficients to be estimated and u_i is a random error term.

The selection issue is considered by allowing the error terms u_i and v_i to be correlated so that:

$$v_i \sim N(0, 1)$$

$$u_i \sim N(0, 1)$$

$$Corr(u_i, v_i) = \rho_{w,OV} \tag{3}$$

The correlation term $\rho_{w,OV}$ will capture any omitted factor like unobserved ability that impacts employability and overeducation risk. The greater its magnitude, the larger the weight of shared unobservables.

The model is estimated by maximum likelihood. We consider all the explanatory variables introduced before (*SocDem, Univ, AddStu, English, PregradExp, PregradMob, PostgradExp, PostgradMob, JSStrat*, *JobRej* and *RegFE*) in both equations. To minimize omitted variable bias, scholars recommend including a wide set of controls for worker heterogeneity (McGuinness, 2006). For identification, it is convenient that Z_i contains at least one variable not included in X_i (exclusion restriction) (Puhani, 2000). We use the regional *employment* rate that corresponds to the respondent's gender as our exclusion restriction (*EMP.RATE*). This variable is retrieved from the Spanish National Statistics Institute for the year 2014 and complemented by Eurostat for the low share of graduates that live abroad (5.53% of the sample).[16] Büchel and van Ham (2003) document that regional unemployment rate does not impact overeducation but affects employment probabilities. Similar

[16] We use the employment rate for the 25–54 age interval, which corresponds with the age range in our sample. Unfortunately, we could not specify separate employment rates for the below and over 30 cohorts due to data unavailability.

Table 3 Average differences in overeducation likelihood between females and males using Inverse probability weighting regression adjustment (IPWRA)

	IPWRA	n	Females	Males
All	0.006 (0.007)	21,074	12,446	8628
Less 30	0.013 (0.009)	12,206	8022	4184
Over 30	0.006 (0.012)	8868	4424	4444

Robust standard errors in parentheses

The reported estimates refer to the average difference in conditional-on-covariates overeducation prevalence between males and females after reweighting on the labour market participation propensity scores

***p < 0.01, **p < 0.05, *p < 0.1

findings are reported in Charalambidou and McIntosh (2021). Moreover, Jauhiainen (2011) uses the unemployment rate as the exclusion restriction in a similar Heckman probit regression for modelling overeducation likelihood. We, instead, use the *employment* rate for such purpose. The rationale is that with an increasing rate of employed people over the working age population, the probability of being working increases, plausibly through lower inactivity rates. At the same time, the extent to which available labour resources are being used is predicted to be unrelated to the risk of overeducation since it does not hinder (at least directly, given other controls) the possibility of finding a matched job.[17] Auxiliary regressions (Table A4 in the Supplementary Material) support that *EMP.RATE* is strongly correlated with employability but uncorrelated with overeducation conditional on covariates. Therefore, this variable can be considered a valid exclusion restriction.

5 Results

5.1 Gender differences in overeducation prevalence based on matching

Table 3 presents the average difference in overeducation risk between females and males using inverse probability weighting regression adjustment, for the whole sample and separately for graduates under and over 30 years of age. That is, the gender average gap in overeducation prevalence is computed while reweighting observations by the employment propensity scores after PSM. Tables A5 and A6 in Supplementary Material report the number of observations per propensity block. Plots for overlapping in the propensity scores by employment status are shown in Figures A5 and A6, Supplementary Material. These diagnostic analyses suggest that covariates are sufficiently balanced, thereby allowing an appropriate comparison.

[17] Whereas unemployment rates are highly correlated with the business cycle so that they could be partially correlated with overeducation prevalence, the cross-sectional variability in employment rates across gender, age and region is likely to stem from differences in activity rates, being far less dependent on the business cycle (see Figures A1 and A2, Supplementary Material).

We find no significant differences in the probability of being overeducated between males and females conditional on their characteristics and their employment propensities. This is sustained for the pooled sample and for the two age cohorts considered. Therefore, it seems the gender difference in overeducation detected in descriptive statistics could be the result of differences in the sample composition of workers (self-selection into the labour market) and job matches based on characteristics. In the following subsection, we move to the Heckman probit regression with interactions to inspect whether the predictors of overeducation exert significantly distinct effects by gender.

5.2 The drivers of overeducation

Tables 4 and 5 report the parameter estimates and robust standard errors for the Heckman probit model with interaction terms specified in Columns (1–3) for the less than 30 and over 30 age groups. Since the specification contains many interaction terms and the coefficients do not have a direct interpretation, we also present the average marginal effects (AME) for each equation. We focus first on the overall results for the drivers of overeducation; then we come back to gender differences.

Starting with the selection equation, the likelihood of being employed is not significantly different by gender for the over 30 cohort, everything else being equal. However, females under 30 are significantly more likely to be working (+5.4%). For both cohorts, labour market participation is positively associated with the employment rate at the place of residence and graduating from health sciences (relative to social sciences). In contrast, graduates from sciences and arts are significantly less likely to work. Labour participation is positively associated with holding a certificate in English, the length of the university studies, pre-graduation mobility through an Erasmus programme, and post-graduation mobility in terms of moving to another region. Similarly, the working likelihood is higher among those who worked while studying, those with two or more years of labour experience, and those who have rejected a job in the past. Conversely, having lived abroad after graduation, the number of past employers and currently studying are negatively associated with labour participation.

For the less than 30 cohort, employability is also more prevalent among those who earned an excellence grant (+3.5%) and graduates in engineering (+5.1%). On the contrary, individuals who accessed their first job through a temporary employment agency (−3.6%) or through the intermediation of the university (−1.4%) are less likely to be working. The same applies to having received a general grant for completing their university studies (−1.3%).

Moving to the outcome equation, there are no gender differences in the prevalence of overeducation for both cohorts, *ceteris paribus*. This is consistent with the results from the matching estimators presented in Table 3. Males and females appear not to have different probabilities of being overeducated conditional on their characteristics, in line with Acosta-Ballesteros et al. (2018). Nonetheless, we explore this in detail in the following subsection. We find that graduates from public universities

Table 4 SE-Probit model parameter estimates and AME for less than 30 cohort

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
FEMALE	−0.334	(0.209)	0.054***	(0.019)	0.027	(0.224)	0.005	(0.010)
SPANISH	−0.011	(0.115)	0.043*	(0.026)	−0.312	(0.538)	−0.123	(0.168)
SPANISH × FEMALE	0.246**	(0.102)			−0.161	(0.114)		
PUB_UNI	−0.054	(0.063)	0.012	(0.010)	0.028	(0.068)	0.025**	(0.011)
PUB_UNI × FEMALE	0.147*	(0.077)			0.093	(0.086)		
SCIENCES	−0.237***	(0.079)	−0.033***	(0.012)	−0.271***	(0.101)	−0.066***	(0.013)
SCIENCES × FEMALE	0.168*	(0.096)			0.032	(0.117)		
ARTS	−0.309***	(0.085)	−0.066***	(0.014)	0.298***	(0.101)	0.060***	(0.016)
ARTS × FEMALE	0.106	(0.102)			−0.123	(0.117)		
ENGINEERING	0.250***	(0.056)	0.051***	(0.010)	−0.365***	(0.061)	−0.091***	(0.011)
ENGINEERING × FEMALE	−0.063	(0.082)			0.011	(0.086)		
HEALTH	0.143*	(0.076)	0.029***	(0.009)	−1.113***	(0.110)	−0.250***	(0.007)
HEALTH × FEMALE	−0.044	(0.086)			−0.134	(0.124)		
GRANT_STU	−0.012	(0.045)	−0.013*	(0.007)	0.131***	(0.049)	0.043***	(0.008)
GRANT_STU × FEMALE	−0.054	(0.054)			0.039	(0.060)		
GRANT_EXC	0.125	(0.078)	0.035***	(0.012)	−0.244***	(0.088)	−0.046***	(0.014)
GRANT_EXC × FEMALE	0.022	(0.099)			0.095	(0.111)		
STU_LENGTH	0.081***	(0.024)	0.025***	(0.004)	−0.116***	(0.027)	−0.025***	(0.004)
STU_LENGTH × FEMALE	0.022	(0.029)			0.032	(0.032)		
MASTER	0.016	(0.046)	0.003	(0.007)	−0.051	(0.051)	−0.027***	(0.008)

Table 4 (continued)

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
MASTER × FEMALE	−0.009	(0.057)			−0.071	(0.064)		
VOC_TRAINING	−0.042	(0.080)	−0.016	(0.011)	0.420***	(0.084)	0.120***	(0.015)
VOC_TRAINING × FEMALE	−0.028	(0.094)			−0.028	(0.102)		
CONT_STUDY	−0.278***	(0.048)	−0.069***	(0.008)	−0.132**	(0.065)	−0.051***	(0.009)
CONT_STUDY × FEMALE	0.036	(0.058)			−0.069	(0.073)		
OLS_CERTIFICATE	0.091*	(0.053)	0.023***	(0.007)	−0.090	(0.059)	−0.038***	(0.009)
OLS_CERTIFICATE × FEMALE	−0.001	(0.063)			−0.082	(0.071)		
CAMB_CERTIFICATE	−0.009	(0.045)	0.020***	(0.007)	−0.209***	(0.051)	−0.071***	(0.008)
CAMB_CERTIFICATE × FEMALE	0.129**	(0.055)			−0.080	(0.064)		
WORK_STU	0.011	(0.043)	0.014**	(0.007)	0.127***	(0.048)	0.024***	(0.008)
WORK_STU × FEMALE	0.063	(0.053)			−0.061	(0.059)		
CURRIC_PRACT	0.051	(0.047)	0.013	(0.008)	−0.030	(0.052)	−0.008	(0.009)
CURRIC_PRACT × FEMALE	−0.003	(0.061)			−7.0e−05	(0.067)		
NONCURRIC_PRACT	−0.003	(0.044)	0.005	(0.007)	−0.119**	(0.049)	−0.019**	(0.008)
NONCURRIC_PRACT × FEMALE	0.031	(0.054)			0.071	(0.060)		
NAT_STAY	−0.041	(0.078)	0.002	(0.013)	−0.047	(0.090)	−0.009	(0.015)
NAT_STAY × FEMALE	0.074	(0.100)			0.023	(0.114)		
ERASMUS	0.157**	(0.073)	0.018*	(0.011)	−0.302***	(0.082)	−0.043***	(0.012)
ERASMUS × FEMALE	−0.124	(0.091)			0.202**	(0.102)		
EXP_2YEARS	0.752***	(0.044)	0.223***	(0.008)	−0.268***	(0.084)	−0.062***	(0.009)
EXP_2YEARS × FEMALE	0.034	(0.054)			0.021	(0.064)		

Table 4 (continued)

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
N_EMPLOYERS	−0.049***	(0.011)	−0.009***	(0.002)	0.053***	(0.013)	0.012***	(0.002)
N_EMPLOYERS × FEMALE	0.020	(0.013)			−0.014	(0.015)		
LAB_MOB	0.201***	(0.056)	0.021**	(0.008)	−0.177***	(0.062)	−0.070***	(0.009)
LAB_MOB × FEMALE	−0.172**	(0.069)			−0.136*	(0.078)		
ABROAD_RESIDENCE	−0.191***	(0.052)	−0.057***	(0.009)	−0.145**	(0.064)	−0.025**	(0.010)
ABROAD_RESIDENCE × FEMALE	−0.027	(0.064)			0.091	(0.077)		
JS_ADV	0.085*	(0.048)	0.005	(0.007)	0.013	(0.052)	0.034***	(0.009)
JS_ADV × FEMALE	−0.098*	(0.059)			0.163**	(0.064)		
JS_TEA	−0.237***	(0.074)	−0.036***	(0.011)	0.283***	(0.089)	0.079***	(0.015)
JS_TEA × FEMALE	0.154*	(0.089)			−0.013	(0.104)		
JS_UNI	−0.050	(0.053)	−0.014*	(0.008)	−0.002	(0.061)	−0.020**	(0.009)
JS_UNI × FEMALE	−0.007	(0.064)			−0.101	(0.074)		
JS_NETWORKS	−0.029	(0.043)	−0.005	(0.006)	0.160***	(0.048)	0.036***	(0.008)
JS_NETWORKS × FEMALE	0.016	(0.053)			−0.045	(0.058)		
REJECT_JOB	0.120***	(0.044)	0.024***	(0.007)	−0.206***	(0.049)	−0.045***	(0.008)
REJECT_JOB × FEMALE	−0.042	(0.055)			0.057	(0.061)		
EMPL.RATE	0.016***	(0.006)	0.004***	(0.002)				
Region fixed effects	YES				YES			
Constant	−0.968**	(0.423)			0.516	(0.575)		
$\rho_{W,OV}$	−0.092 (0.205)							
Pseudo R squared	0.120				0.143			

Table 4 (continued)

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
Total observations	15,667							
Selected	12,206							
Non-selected	3471							

***p < 0.01, **p < 0.05, *p < 0.1

Table 5 SE-Probit model parameter estimates and AME for over 30 cohort

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
FEMALE	−0.270	(0.282)	0.013	(0.020)	0.048	(0.287)	−0.004	(0.014)
SPANISH	−0.114	(0.169)	0.017	(0.032)	−0.436	(0.377)	−0.184	(0.118)
SPANISH × FEMALE	0.349**	(0.162)			−0.268	(0.176)		
PUB_UNI	0.092	(0.065)	0.002	(0.011)	0.332***	(0.066)	0.115***	(0.015)
PUB_UNI × FEMALE	−0.155*	(0.091)			0.056	(0.095)		
SCIENCES	−0.196**	(0.083)	−0.024*	(0.012)	−0.076	(0.090)	−0.008	(0.018)
SCIENCES × FEMALE	0.166	(0.108)			0.087	(0.113)		
ARTS	−0.098	(0.083)	−0.026**	(0.012)	0.072	(0.082)	0.018	(0.017)
ARTS × FEMALE	−0.034	(0.109)			−0.042	(0.109)		
ENGINEERING	0.075	(0.053)	0.005	(0.010)	−0.364***	(0.052)	−0.093***	(0.013)
ENGINEERING × FEMALE	−0.100	(0.082)			0.134	(0.082)		
HEALTH	0.356***	(0.123)	0.060***	(0.016)	−0.655***	(0.124)	−0.245***	(0.022)
HEALTH × FEMALE	−0.170	(0.142)			−0.197	(0.141)		
GRANT_STU	−0.004	(0.053)	−0.006	(0.008)	0.219***	(0.052)	0.067***	(0.011)
GRANT_STU × FEMALE	−0.042	(0.070)			−0.019	(0.071)		
GRANT_EXC	0.064	(0.110)	0.006	(0.018)	0.011	(0.101)	−0.057**	(0.027)
GRANT_EXC × FEMALE	−0.072	(0.157)			−0.363**	(0.168)		
STU_LENGTH	0.047*	(0.025)	0.012***	(0.004)	−0.082***	(0.024)	−0.026***	(0.005)
STU_LENGTH × FEMALE	0.008	(0.034)			0.003	(0.033)		
MASTER	0.022	(0.052)	0.003	(0.008)	−0.107**	(0.051)	−0.057***	(0.011)
MASTER × FEMALE	−0.014	(0.071)			−0.135*	(0.072)		
VOC_TRAINING	0.012	(0.059)	−0.007	(0.009)	0.201***	(0.055)	0.083***	(0.012)

Table 5 (continued)

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
VOC_TRAINING × FEMALE	−0.078	(0.078)			0.109	(0.078)		
CONT_STUDY	−0.193***	(0.054)	−0.033***	(0.008)	−0.130**	(0.060)	−0.029**	(0.012)
CONT_STUDY × FEMALE	0.082	(0.075)			0.055	(0.079)		
OLS_CERTIFICATE	0.241***	(0.065)	0.042***	(0.009)	−0.048	(0.065)	−0.033***	(0.012)
OLS_CERTIFICATE × FEMALE	−0.101	(0.083)			−0.088	(0.078)		
CAMB_CERTIFICATE	−0.032	(0.053)	0.001	(0.009)	−0.086	(0.054)	−0.027**	(0.012)
CAMB_CERTIFICATE × FEMALE	0.069	(0.075)			0.005	(0.076)		
WORK_STU	0.169***	(0.051)	0.038***	(0.008)	−0.036	(0.058)	−0.003	(0.012)
WORK_STU × FEMALE	−0.004	(0.069)			0.068	(0.073)		
CURRIC_PRACT	−0.011	(0.047)	0.009	(0.008)	−0.081*	(0.045)	−0.013	(0.010)
CURRIC_PRACT × FEMALE	0.093	(0.066)			0.080	(0.066)		
NONCURRIC_PRACT	−0.081	(0.050)	−0.027***	(0.008)	−0.029	(0.050)	0.005	(0.011)
NONCURRIC_PRACT × FEMALE	−0.071	(0.067)			0.072	(0.071)		
NAT_STAY	0.133**	(0.067)	0.021**	(0.011)	−0.071	(0.063)	−0.036**	(0.014)
NAT_STAY × FEMALE	−0.073	(0.093)			−0.072	(0.088)		
ERASMUS	0.128	(0.097)	0.026*	(0.015)	−0.086	(0.092)	−0.006	(0.021)
ERASMUS × FEMALE	−0.020	(0.132)			0.138	(0.130)		
EXP_2YEARS	1.117***	(0.050)	0.266***	(0.007)	−0.141	(0.210)	−0.056***	(0.021)
EXP_2YEARS × FEMALE	0.113*	(0.068)			0.061	(0.100)		
N_EMPLOYERS	−0.083***	(0.012)	−0.016***	(0.002)	0.035*	(0.019)	0.008***	(0.003)
N_EMPLOYERS × FEMALE	0.021	(0.016)			−0.026	(0.018)		

Table 5 (continued)

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
LAB_MOB	0.076	(0.066)	0.027**	(0.011)	−0.114*	(0.065)	−0.044***	(0.015)
LAB_MOB × FEMALE	0.083	(0.094)			−0.034	(0.093)		
ABROAD_RESIDENCE	−0.286***	(0.066)	−0.040***	(0.011)	−0.150*	(0.082)	−0.023	(0.017)
ABROAD_RESIDENCE × FEMALE	0.200**	(0.096)			0.127	(0.110)		
JS_ADV	0.035	(0.053)	−0.009	(0.008)	−0.008	(0.050)	0.011	(0.011)
JS_ADV × FEMALE	−0.138*	(0.071)			0.078	(0.073)		
JS_TEA	−0.070	(0.079)	−0.018	(0.012)	0.253***	(0.083)	0.103***	(0.017)
JS_TEA × FEMALE	−0.014	(0.106)			0.125	(0.111)		
JS_UNI	−0.152**	(0.059)	−0.011	(0.009)	−0.085	(0.065)	−0.019	(0.013)
JS_UNI × FEMALE	0.190**	(0.079)			0.040	(0.086)		
JS_NETWORKS	−0.027	(0.046)	−0.012	(0.007)	0.159***	(0.045)	0.033***	(0.010)
JS_NETWORKS × FEMALE	−0.045	(0.064)			−0.110*	(0.064)		
REJECT_JOB	0.204***	(0.051)	0.034***	(0.008)	−0.215***	(0.055)	−0.056***	(0.011)
REJECT_JOB × FEMALE	−0.096	(0.071)			0.092	(0.070)		
EMPL.RATE	0.017**	(0.008)	0.004**	(0.002)				
Region fixed effects	YES				YES			
Constant	−1.214**	(0.576)			0.234	(0.517)		
$\rho_{W,OV}$	0.152 (0.402)							
Pseudo R squared	0.192				0.075			
Total observations	11,140							
Selected	8868							

Table 5 (continued)

	Selection equation Prob (*WORK* = 1)				Outcome equation *OVEREDUC*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficient	Robust SE	AME	Robust SE	Coefficient	Robust SE	AME	Robust SE
Non-selected	2272							

***p < 0.01, **p < 0.05, *p < 0.1

are more likely to be overeducated (+2.5% for less than 30 and +11.5% for over 30, respectively). In contrast, overeducation is unrelated to being Spanish.

Overeducation is less prevalent among graduates in the fields of health (−25% and −24.5%), engineering (−9.1% and −9.3%) and sciences (−6.6%, only for less than 30 cohort). However, graduates in the arts are the most likely to be overeducated (+6%, only for less than 30 cohort). This is consistent with Aina and Pastore (2020), Barone and Ortiz (2011), Meroni and Vera-Toscano (2017) and Turmo-Garuz et al. (2019). Something similar happens with those who received a grant for completing their university studies (+4.3% and +6.7%) and those who have a vocational training certificate (+12% and +8.3%). The former can reflect a lower family background (Verhaest & Omey, 2010) whereas the latter might imply that, even though they hold a university degree, those with additional vocational training studies find it easier to work in occupations that only demand such formation. Having worked while studying also increases overeducation prevalence, but only for graduates under 30 (+2.4%). Conversely, the likelihood of overeducation is negatively related with holding a Master's (−2.7% and −5.7%), the length of tertiary studies (−2.5% and −2.6% per year), currently studying (−5.1% and −2.9%) and having earned an excellence grant (−4.6% and −5.7%), which can be understood as a proxy of innate ability. This confirms earlier results by Devillanova (2013), Di Pietro and Urwin (2006), Verhaest and Omey (2010), and suggests that the greater the accumulation of human capital, the lower the likelihood of overeducation. We also find that having a certificate in English, either from the OLS (−3.8% and −3.3%) or from Cambridge (−7.1% and −2.7%), significantly protects Spanish graduates from overeducation, in line with Albert and Davía (2018).

Concerning pre-graduation mobility, having participated in the Erasmus programme reduces the probability of being overeducated, but only for graduates under 30 (−4.3%). Curiously, completing part of their university studies at another Spanish university reduces overeducation only for the over 30 cohort (−3.6%). As discussed before, this could be explained by the acquisition of soft skills and personal independence during the stays. The same happens with having moved to another Spanish region to look for a job (−7% and −4.4%). Therefore, national mobility helps to reduce the incidence of educational mismatch, in line with Devillanova (2013). Similarly, we do not find evidence that internships within the study programme (curricular practices) reduce the likelihood of being overeducated. This could indicate that internship programmes offered by Spanish universities are not effective at providing graduates with the necessary on-the-job training to get matched jobs. However, non-curricular practices do reduce the likelihood of overeducation for the less than 30 cohort (−2%). Having at least two years of work experience also reduces the incidence of education mismatch (−6.2% and −5.6%), in line with Kiker et al. (1997) and Robst (2008).

We also document that those with a larger number of past employers are more likely to be overeducated in their jobs (+1.2% and +0.8%). This contrasts with Romanov et al. (2017) and Rubb (2013), who suggest that firm-switching reduces the likelihood of overeducation. Our findings could be explained by the particularities of our study period. Borgna et al. (2019) show that in the aftermath of the 2008 economic crisis, overeducation risks were higher among those who experienced job

mobility after the outbreak of the crisis. Regarding job search strategies, looking for the first job through advertisements on the internet or in newspapers (+3.4%, only for the less than 30 cohort), temporal employment agencies (+7.9% and +10.3%) or work/family networks (+3.6% and 3.3%) is positively related to overeducation prevalence. This is consistent with the findings by Albert and Davía (2018) using the same dataset. Please note that the omitted category gathers other options like self-employment or becoming a civil servant, which seem to be the most effective ways to avoid overeducation at the beginning of the labour career. Finally, having rejected a job in the past because it was considered unsuitable is negatively related to the likelihood of overeducation (−4.5% and −5.6%). This could account for individuals from less advantaged backgrounds being in more need of work so that they are more forced to accept jobs requiring lower qualifications (Barone & Ortiz, 2011). Another explanation relates to McCormick's theory of signalling during the job search (McCormick, 1990). Since individuals know their own ability, those who are more skilled might be reluctant to accept overeducated job positions and will wait for a better job opportunity (they have a higher reserve wage) to signal their capabilities.

The error terms of the two equations are not significantly correlated. This implies that there seems not to be common unobservables in the residuals of the two equations given the wide set of observed characteristics. Robustness checks on this are discussed in Sect. 5.4.

To rule out potential concerns and misinterpretation of our results, we want to clarify the following. Firstly, the interaction coefficients do not have an associated AME. The reason is that the interaction term is contained in the AME formula for each variable. Furthermore, Greene (2010) warns about the difficulties of interpreting the partial effects for interaction terms in non-linear models. Secondly, as a direct consequence of this, the significance of the AME does not always correspond to that of the associated coefficient estimates. This is because if there is some heterogeneity in the effect of a given covariate on the dependent variable by gender, the effects might cancel out and lead to an AME non-distinguishable from zero. This is also affected by the coefficient covariances. Thirdly, the statistical significance (and the sign) of the interaction effects based on the t-test must be taken with caution (Ai & Norton, 2003).[18]

5.3 Marginal effects and gender differences in predictors

To further examine the existence of gender differences in the drivers of overeducation, we proceed as follows. First, we compute the overall AME for the outcome equation separately for males and females for each age cohort. Then, we test the hypothesis of whether they are statistically different (i.e. H0: $AME_{female} = AME_{male}$) using a Wald test. In this way, computation of the separate AME considers the differences in characteristics and coefficient estimates across gender, with the additional advantage that the coefficients are measured on the same scale and are directly

[18] This is because, unlike the linear case, the interaction effect in a non-linear model could be positive or negative depending on the values of all the variables that enter the standard normal cumulative density function.

Table 6 AME by gender and age cohort

Variable	Less than 30				Over 30			
	Males	Females	Difference	Wald test	Males	Females	Difference	Wald test
SPANISH	−0.090	−0.139	−0.083	1.92	−0.133	−0.232	−0.099	3.32*
PUB_UNI	0.007	0.034**	0.027	1.43	0.101***	0.127***	0.025	0.78
SCIENCES	−0.070***	−0.064***	0.006	0.05	−0.020	0.004	0.024	0.49
ARTS	0.082***	0.047**	−0.034	0.96	0.023	0.012	−0.011	0.12
ENGINEERING	−0.093***	−0.089***	0.003	0.02	−0.113***	−0.073***	0.039	2.39
HEALTH	−0.209***	−0.270***	−0.061	17.2***	−0.208***	−0.279***	−0.071	2.67
GRANT_STU	0.035***	0.047***	0.011	0.48	0.067***	0.065***	−0.002	0.01
GRANT_EXC	−0.060**	−0.038**	0.021	0.64	0.002	−0.114***	−0.116	4.74**
STU_LENGTH	−0.030***	−0.022***	0.008	0.83	−0.026***	−0.026***	4.4e−04	0.01
MASTER	−0.013	−0.033***	−0.020	1.37	−0.033**	−0.078	−0.045	3.92**
VOC_TRAINING	0.125***	0.117***	−0.008	0.07	0.062***	0.102	0.039	2.76*
CONT_STUDY	−0.037**	−0.057***	−0.019	1.08	−0.037**	−0.022	0.015	0.37
OLS_CERTIFICATE	−0.022	−0.046***	−0.023	1.55	−0.018	−0.046***	−0.027	1.29
CAMB_CERTIFICATE	−0.055***	−0.078***	−0.022	1.85	−0.026	−0.026	−7.7e−04	0.01
WORK_STU	0.034***	0.019**	−0.014	0.87	−0.013	0.007	0.020	0.81
CURRIC_PRACT	−0.007	−0.007	3.0e−04	0.01	−0.025*	−0.001	0.023	1.24
NONCURRIC_PRACT	−0.031**	0.013	0.018	1.36	−0.007	0.017	0.024	1.28
NAT_STAY	−0.012	−0.006	0.006	0.05	−0.024	−0.047**	−0.023	0.70
ERASMUS	−0.073***	−0.026*	0.047	3.66*	−0.028	0.014	0.043	1.11
EXP_2YEARS	−0.065***	−0.060***	0.005	0.07	−0.061**	−0.049*	0.011	0.18
N_EMPLOYERS	0.013***	0.010***	−0.003	0.65	0.012***	0.004	−0.008	2.17
LAB_MOB	−0.044***	−0.082***	−0.038	4.04**	−0.036*	−0.050**	−0.014	0.25
ABROAD_RESIDENCE	−0.040**	0.017	0.022	1.27	−0.041*	−0.005	0.035	1.13
JS_ADV	0.004	0.049***	0.045	6.75***	−0.003	0.024	0.027	1.54

Table 6 (continued)

Variable	Less than 30				Over 30			
	Males	Females	Difference	Wald test	Males	Females	Difference	Wald test
JS_TEA	0.079***	0.078***	−7.0e−04	0.01	0.079***	0.124***	0.044	1.65
JS_UNI	−0.001	−0.029**	−0.028	2.00	−0.023	−0.015	0.008	0.11
JS_NETWORKS	0.043***	0.032***	−0.011	0.47	0.049***	0.017	−0.032	2.62
REJECT_JOB	−0.053***	−0.040***	0.013	0.70	−0.0700***	−0.041**	0.028	1.64

***p < 0.01, **p < 0.05, *p < 0.1

comparable. Columns 1–2 (5–6) in Table 6 presents the AME for males and females under 30 (over 30). Column 3 (7) reports the difference between the two and column 4 (8) the chi-squared statistic for the significance of the difference.

We find that graduating from health sciences protects females from overeducation to a greater extent than males for the less than 30 cohort. For graduates over 30, the difference is not statistically significant. As regards other fields of study, males from architecture and engineering degrees seem to be less likely to be overeducated than females for both cohorts. However, the difference is not significant. Descriptive statistics (see Table 1) show that health sciences is a female-dominated field of knowledge while males are a majority in engineering and architecture. Statistical discrimination might make employers believe males (females) are better at engineering (health sciences). Independently of whether females are initially, *on average*, less productive in STEM fields of knowledge due to past discrimination or lower self-esteem (Gneezy et al., 2003), implicit discrimination might be present. Social stereotypes as presented in Bordalo et al. (2019) might unconsciously make employers believe that females are better for health care and males for mechanics.

Strikingly, having moved to another Spanish region protects females more than males from overeducation in the less than 30 cohort. This could be associated with the theory of differential overqualification developed by Frank (1978) and the findings of Redmond and McGuinness (2019). If females are more reluctant to move and stick more to local markets, the fact of having moved to another region might signal the possession of soft skills that could be valued more in the case of females. Nonetheless, there are no differences for the over 30 age group. The estimates also indicate that holding a Master's degree signals males at a greater magnitude than females in the over 30 cohort. Additionally, Erasmus (national) stays reduce the incidence of overeducation only for males (females) in the less than 30 cohort.

Remarkably, the increased prevalence of overeducation associated with having looked for the first job through advertisements is only significant for females in the less than 30 cohort. This suggests that accessing first employment though public calls in which there are no intermediaries seems to damage females. In job interviews with more candidates, females might shy away from showing their capacities and end up working in overeducated jobs in a greater proportion than males. For the rest of the variables, although the AME slightly differ in magnitude, the differences are not statistically different. Overall, human capital seems to be valued equally for males and females.

5.4 Robustness checks

We have performed some robustness checks. Firstly, we inspected collinearity problems. The Variance Inflation Factor (VIF) after separate OLS regressions with interactions are 5.50 and 6.33 (5.12 and 4.98) for the selection and outcome equations for the less than 30 (over 30) cohort.[19] All of them lie within acceptable boundaries.

[19] A large share of the correlation is due to the inclusion of interaction terms. VIF values after OLS regressions without interactions shrink to 2.89 and 4.26 (2.75 and 2.48) for the selection and outcome equations for the less than 30 (over 30) cohort.

Secondly, we ran separate Heckman probit regressions for males and females (Tables A7–A8, Supplementary Material). Although the coefficient estimates cannot be compared due to scale differences, the statistical significance of the variables is consistent with our main findings.

Thirdly, we inspected the robustness of the non-significance of the gender dummy in the overeducation equation and the correlation between the error terms of the selection and outcome equations by conducting a stepwise estimation in which each group of variables was included in the model specification sequentially. Figures A3 and A4 in Supplementary Material present the coefficient estimates for the gender dummy and the rho parameter ($\rho_{W,OV}$). We document that the coefficient estimate for the gender dummy is consistently non-significant under different specifications. Even though we face the risk of omitted variables, under the standard assumption that selection in unobservables is proportional to observables, the stability in the female dummy across specification indicates unobserved factors should weight more than observables to rule out the non-significant effect. Additionally, the correlation between the error terms of the selection and the outcome equations becomes non-significant after the inclusion of the regional fixed effects, which might capture omitted factors at the local labour market level. In this regard, we further examined the sensitivity of our findings to different potential values of the correlation between the error terms using the procedure recently developed by Cook et al. (2021). Our main findings hold (see Figure A7 in Supplementary Material).

Fourthly, we examined the potential existence of non-random measurement error in our self-assessed measure of overeducation that could vary by gender or correlate with workers' ability. To this end, we first implemented the test to detect heteroskedasticity in binary outcomes as proposed by Wooldridge (2010). Next, we ran a heteroskedastic probit regression in which the variance of the error term is assumed to be an exponential function of gender and having earned an excellence grant as a crude proxy of ability. The results and technical details are shown in Appendix E, Table A9. According to this check, there is no evidence of measurement error associated with gender or ability. Therefore, even though we cannot completely rule out the possibility of measurement error in overeducation because of being self-reported, it appears this does not alter our findings.

Fifthly, instead of the dummy variable, we used an ordinal indicator capturing the *degree* of overeducation. Details on the variable construction and the parameter estimates are presented in the Supplementary Material (Tables A1 and A2). The significance and direction of the effects remain unchanged. Finally, we repeated our analysis measuring overeducation using (i) the statistical approach (SA), and (ii) a novel method that combines the earnings premium procedure proposed by Gottschalk and Hansen (2003) with the statistical approach. The variable construction and the regression results are shown in Sections F and G in Supplementary Material. Although the prevalence of overeducation slightly differs across the three measures (which is something common in related studies, see Verhaest and Omey (2010)), in all cases the percentage of overeducated workers is larger among females and for the over 30 cohort.[20]

[20] Considering all the sample, the share of overeducated workers according to the ISA, SA and earnings methods are 25.9%, 21.9% and 16.3%, respectively (Table A18 in Supplementary Material).

Additionally, overeducation based on the ISA approach is consistent with other indicators of a bad job match like the share of graduates that regret having studied the attained college degree or having even enrolled at university studies (Table A19 in Supplementary Material). Importantly, the results obtained from regression analyses using these alternative measures of overeducation are similar to those derived from the self-assessed approach.[21]

6 Conclusions

This paper studied the drivers of overeducation of recent university graduates paying particular attention to the existence of gender differences. Using a representative sample of two age groups from the same graduation cohort in Spain observed four years after the completion of their university studies, we first examined gender differences in overeducation prevalence using matching estimators. Next, we estimated a Heckman probit model with interactions between the explanatory variables and the gender dummy that considers self-selection into the labour market.

Our results show that, conditional on their characteristics, males and females have about the same probability of being overeducated in their jobs. The differences in overeducation prevalence documented in descriptive statistics appear to be driven by differences in the probability of being employed and gender differences in some of the drivers of overeducation. Conditional on labour market participation, the probability of being overeducated is lower among graduates from sciences, engineering and health sciences, among those who earned an excellence grant, and among those with an English language certificate. Mobility, labour experience and years of tertiary schooling also reduce the prevalence of overeducation. In contrast, graduates from arts, those who worked while studying, those who accessed their first employment through advertisements, temporary employment agencies or family networks, and those who benefited from a scholarship (proxy for lower family background) are more exposed to overeducation. Interestingly, we document some differences in the magnitude of the effects across age cohorts, with non-curricular practices or Erasmus stays reducing overeducation only among graduates aged less than 30.

By analysing the differences in the drivers of overeducation by gender, the study has assessed potential implicit discrimination based on characteristics. Based on separate regressions by age cohort with gender interactions, we exploit the non-linearity of the model and examine the average marginal effects by gender. Apart from the finding that there are no gender differences everything else being equal, personal characteristics appear to have the same effect on the likelihood of overeducation for

[21] In the case of the overeducation based on the earnings penalty method (Tables A16 and A17, Supplementary Material), it appears males and females have different overeducation risks. Notwithstanding this, these estimates need to be taken with care. First, they are based on estimated wages, which involve some uncertainty around the true values. Second, we face a potential non-random missing data problem since we lack information about wages for some graduates that are working.

both sexes. As such, the characteristics of highly educated males and females are valued equally by the market.

Nevertheless, the estimates show that, *ceteris paribus*, female (male) graduates from health sciences (engineering and architecture) under 30 are less likely to be overeducated in their jobs than their male (female) peers. This could reflect a stereotyping of cognitive ability or statistical discrimination that emanates from the usual higher proportion of males (females) in the STEM (health science) fields of knowledge. Similarly, females are more protected against overeducation than males if they earned an excellence grant (for the over 30 cohort), which could partially capture intelligence or innate ability. Conversely, females under 30 are significantly more likely to be overeducated if they accessed their first job through public advertisements that involved directly contacting the employer. It seems that when third parties intermediate, the gender gap vanishes. We also show that moving to another Spanish region for labour-related reasons contributes to reduce the likelihood of overeducation to a greater extent among females. This could be due to females' higher reluctance to move, which could convey a positive signal for female movers.

From a policy viewpoint, our results contribute to the debate about the economic consequences of the educational mismatch phenomenon. Education represents one of the most relevant public expenditures and its return to society is only achieved when graduates are well matched to their jobs. The evidence on the differences in overeducation by gender based on the field of study suggests that educational institutions should redesign the supply side of their study programmes and incentivize females (males) to access masculinized (feminized) fields of study to avoid stereotypes and statistical discrimination. Although our findings show, at least for the case of recent university graduates, that females and males have an equal likelihood of being overeducated conditional on characteristics, the fact that labour mobility and the completion of national stays reduces the risk of overeducation to a greater extent among females deserves further attention. Since females have traditionally been more reluctant to move, it appears that moving to another region to look for a job is an effective job search strategy for them to avoid the risk of overeducation. This calls for educational institutions to encourage graduates, and particularly females, to undertake stays at other Spanish universities, since apart from increasing their chances of finding a suitable job, it also provides them with soft skills that are highly valued by the market. On a more general level, public policies aimed at improving the matching of university graduates with local job positions appear to be necessary in Spain to reduce the share of university graduates that end up in overeducated positions.

Our study has some limitations. The cross-sectional nature of the dataset limits the analysis to a 'between' comparison at a specific period. Therefore, since we cannot completely rule out omitted confounders, our findings need to be interpreted with caution. Future research using longitudinal data could extend our study by examining the dynamics of gender differences in overeducation throughout the labour career or in expansionary rather than recessionary periods. We also lack information on family background or having children. Future studies should go deeper into the effect of parental characteristics and family responsibilities on education mismatch.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40888-022-00270-y.

Acknowledgements We would like to thank three anonymous referees because of all the helpful comments and suggestions received during the editorial review process. The authors declare that the work described has not been published before, that is not under consideration for publication anywhere else, and that its publication has been approved by all co-authors.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Conflict of interest None.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica, 74*, 235–267.

Acosta-Ballesteros, J., Osorno-del Rosal, M. P., & Rodríguez-Rodríguez, O. M. (2018). Overeducation of young workers in Spain: How much does the first job matter? *Social Indicators Research, 138*, 109–139.

Agopsowicz, A., Robinson, C., Stinebrickner, R., & Stinebrickner, T. (2020). Careers and mismatch for college graduates: College and non-college jobs. *Journal of Human Resources, 55*(4), 1194–1221.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters, 80*, 123–129.

Aina, C., & Pastore, F. (2020). Delayed graduation and overeducation in Italy: A test of the Human Capital model versus the screening hypothesis. *Social Indicators Research, 152*, 533–553.

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

Albert, C., & Davia, M. A. (2018). Job search strategies and underemployment in recent graduates first jobs in Spain. *Revista De Economía Aplicada, 78*(26), 21–41.

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, 3*, 434–452.

Astorne-Figari, C., & Speer, J. D. (2019). Are changes of major major changes? The roles of grades, gender, and preferences in college major switching. *Economics of Education Review, 70*, 75–93.

Baert, S., Cockx, B., & Verhaest, D. (2013). Overeducation at the start of the career: Stepping store or trap? *Labour Economics, 25*, 123–140.

Bar-Haim, E., Chauvel, L., & Hartung, A. (2019). More necessary and less sufficient: An age-period-cohort approach to overeducation from a comparative perspective. *Higher Education, 78*, 479–499.

Barnichon, R., & Zylberberg, Y. (2019). Underemployment and the tickle-down of unemployment. *American Economic Journal: Macroeconomics, 11*(2), 40–78.

Barone, C., & Assirelli, G. (2020). Gender segregation in higher education: An empirical test of seven explanations. *Higher Education, 79*, 55–78.

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.

Barron, J. M., Black, D. A., & Loewenstein, M. A. (1989). Job matching and on-the-job-training. *Journal of Labor Economics, 7*(1), 1–19.

Battu, H., & Sloane, P. J. (2004). Over-education and ethnic minorities in Britain. *The Manchester School, 72*(4), 535–559.

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

Beckhusen, J., Florax, R. J. G. M., Poor, J., & Waldorf, B. S. (2013). Attracting global talent and then what? Overeducated immigrants in the United States. *Journal of Regional Science, 53*(5), 834–854.

Belfield, C. (2010). Over-education: What influence does the workplace have? *Economics of Education Review, 29*, 236–245.

Bertrand, M., Chugh, D., & Mullainathan, S. (2005). Implicit discrimination. *American Economic Review, 95*(2), 94–98.

Blanchard, O., & Jimeno, J. F. (1995). Structural unemployment: Spain versus Portugal. *American Economic Review P&p, 85*(2), 212–218.

Bordalo, P., Coffman, K., Gennaioli, N., & Schleifer, A. (2019). Beliefs about gender. *American Economic Review, 109*(3), 739–773.

Borgna, C., Solga, H., & Protsch, P. (2019). Overeducation, labor market dynamics, and economic downturn in Europe. *European Sociological Review, 35*(1), 116–132.

Büchel, F., & van Ham, M. (2003). Overeducation, regional labor markets and spatial flexibility. *Journal of Urban Economics, 53*, 482–493.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness and career choice. *The Quarterly Journal of Economics, 129*(3), 1409–1447.

Capsada-Munsech, Q. (2019). Measuring overeducation: Incidence, correlation and overlaps across indicators and countries. *Social Indicators Research, 145*, 279–301.

Caroleo, F. E., & Pastore, F. (2018). Overeducation at a glance: Determinants and wage effects of the educational mismatch, looking at the AlmaLaurea data. *Social Indicators Research, 137*(3), 999–1032.

Carroll, D., & Tani, M. (2015). Job search as a determinant of graduate over-education: Evidence from Australia. *Education Economics, 23*(5), 631–644.

Cattani, L., Guidetti, G., & Pedrini, G. (2018). Overeducation among Italian graduates: Do different measures diverge? *Economia Politica, 35*, 491–521.

Charalambidou, C., & McIntosh, S. (2021). Over-education in Cyprus: Micro and macro determinants, persistence and state dependence. A dynamic panel analysis. *The Manchester School, 89*, 172–189.

Charlot, O., & Decreuse, B. (2010). Over-education for the rich, under-education for the poor: A search-theoretic microfoundation. *Labour Economics, 17*, 886–896.

Chevalier, A. (2003). Measuring overeducation. *Economica, 70*, 509–531.

Chevalier, A. (2007). Education, occupation and career expectations: Determinants of the gender pay gap for UK graduates. *Oxford Bulletin of Economics and Statistics, 69*(6), 819–842.

Chevalier, A., & Lindley, J. (2009). Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society Series A, 172*(2), 307–337.

Consoli, D., & Sánchez-Barrioluengo, M. (2019). Polarization and the growth of low-skill service jobs in Spanish local labor markets. *Journal of Regional Science, 59*, 145–162.

Cook, J. A., Lee, J. S., & Newberger, N. (2021). On identification and estimation of Heckman models. *Stata Journal, 21*(4), 972–998.

Crossman, J. E., & Clarke, M. (2010). International experience and graduate employability: Stakeholder perceptions on the connection. *Higher Education, 59*(5), 599–613.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics, 132*(4), 1593–1640.

Deming, D. J., & Noray, K. (2020). Earnings dynamics, changing job skills, and STEM careers. *The Quarterly Journal of Economics, 135*(4), 1965–2005.

Deming, D. J., Yuchtman, N., Abulafi, A., Goldin, C., & Katz, L. F. (2016). The value of postsecondary credentials in the labor market: An experimental study. *American Economic Review, 106*(3), 778–806.

Denning, J. T., Marx, B. M., & Turner, L. J. (2018). Propelled: The effects of grants on graduation, earnings and welfare. *American Economic Review: Applied Economics, 11*(3), 193–224.

Devillanova, C. (2013). Over-education and spatial flexibility: New evidence from Italian survey data. *Papers in Regional Science, 92*(3), 445–464.
Di Pietro, G. (2012). Does studying abroad cause international labor mobility? Evidence from Italy. *Economics Letters, 117*, 632–635.
Di Pietro, G., & Urwin, P. (2006). Education and skills mismatch in the Italian graduate labour market. *Applied Economics, 38*, 79–93.
Dolado, J. J., Felgueroso, F., & Jimeno, J. F. (2000). Youth labour markets in Spain: Education, training, and crowding-out. *European Economic Review, 44*, 943–956.
Dolado, J. J., García-Serrano, C., & Jimeno, J. F. (2002). Drawing lessons from the boom of temporary jobs in Spain. *Economic Journal, 112*(480), F270–F295.
Dolado, J. J., Jansen, M., & Jimeno, J. F. (2009). On the job-search in a matching model with heterogeneous jobs and workers. *Economic Journal, 119*, 200–228.
Dolton, P. J., & Silles, M. A. (2008). The effects of over-education on earnings in the graduate labour market. *Economics of Education Review, 27*(2), 125–139.
Duncan, G. J., & Hoffman, S. D. (1981). The incidence of wage effects of overeducation. *Economics of Education Review, 1*(1), 75–86.
Escalonilla, M., Cueto, B., & Pérez-Villadóniga, M. J. (2021). Long-term effects on youth career of entering the labour market during the Great Recession. *Applied Economics, 53*(49), 5643–5657.
Flory, J. A., Leibbrandt, A., & List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *The Review of Economic Studies, 82*(1), 122–155.
Frank, R. (1978). Why women earn less: The theory and estimation of differential over-qualification. *American Economic Review, 68*(3), 360–373.
Frei, C., & Sousa-Poza, A. (2012). Overqualification: Permanent or transitory? *Applied Economics, 44*, 1837–1847.
Gneezy, U., Nierdele, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics, 118*(3), 1049–1074.
Goldin, C., Katz, L. F., & Kuziemko, I. (2006). The homecoming of American College women: The reversal of the College Gender Gap. *Journal of Economic Perspectives, 20*(4), 133–156.
Gottschalk, P., & Hansen, M. (2003). Is the proportion of college workers in noncollege jobs increasing? *Journal of Labor Economics, 21*(2), 449–471.
Green, F., & McIntosh, S. (2007). Is there a genuine under-utilization of skills among the over-qualified? *Applied Economics, 39*, 427–439.
Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters, 107*, 291–296.
Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review, 29*, 911–922.
Guryan, J., & Charles, K. K. (2013). Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *The Economic Journal, 123*(572), F417–F432.
Hartog, J. (2000). Overeducation and earnings: Where are we, where should we go? *Economics of Education Review, 19*(2), 131–147.
Jauhiainen, S. (2011). Overeducation in the Finnish regional labour markets. *Papers in Regional Science, 90*, 573–588.
Jetter, M., & Walker, J. K. (2020). Gender differences in performance and risk-taking among children, teenagers, and college students: Evidence from Jeopardy! *The BE Journal of Economic Analysis and Policy, 20*(2), 20190179.
Jiang, X. (2021). Women in STEM: Ability, preference, and value. *Labour Economics, 70*, 101991.
Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy, 87*(5), 972–990.
Khoo, P., & Ost, B. (2018). The effect of graduating with honors on earnings. *Labour Economics, 55*, 149–162.
Kifle, T., Kler, P., & Fleming, C. M. (2019). The assimilation of Australian immigrants: Does occupation matter? *Applied Economics, 51*(17), 1841–1854.
Kiker, B. F., Santos, M. C., & Mendes de Olivera, M. (1997). Overeducation and undereducation: Evidence from Portugal. *Economics of Education Review, 16*(2), 111–125.
Kleinjans, K. J. (2009). Do gender differences in preferences for competition matter for occupational expectations? *Journal of Economic Psychology, 30*, 701–710.

Krabel, S., & Flöther, C. (2014). Here today, gone tomorrow? Regional labour mobility of German university graduates. *Regional Studies, 48*(10), 1609–1627.
Kratz, F., & Netz, N. (2018). Which mechanisms explain monetary returns to international student mobility? *Studies in Higher Education, 43*(2), 375–400.
Le, A. T., Miller, P. W., Slutske, W. S., & Martin, N. G. (2011). Attitudes towards risk and the gender wage gap. *Labour Economics, 18*, 55–561.
Lenton, P. (2012). Over-education across British Regions. *Regional Studies, 46*(9), 1121–1135.
Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education, 4*, 283–326.
Liu, K., Salvanes, K. G., & Sorensen, E. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. *European Economic Review, 84*, 3–17.
McCormick, B. (1990). A theory of signalling during job search, employment efficiency, and "stigmatised" jobs. *Review of Economic Studies, 57*, 299–313.
McGoldrick, K., & Robst, J. (1996). Gender differences in overeducation: a test of the theory of differential overqualification. *American Economic Review: P&p, 86*(2), 280–284.
McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys, 20*, 387–418.
McGuinness, S., & Bennett, J. (2007). Overeducation in the graduate labour market: A quantile regression approach. *Economics of Education Review, 26*, 521–531.
McGuinness, S., Bergin, A., & Whelan, A. (2018). Overeducation in Europe: Trends, convergence and drivers. *Oxford Economic Papers, 70*(4), 994–1015.
McGuinness, S., Pouliakas, K., & Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys, 32*(4), 985–1015.
McGuinness, S., Whelan, A., & Bergin, A. (2016). Is there a role for higher education institutions in improving the quality of first employment? *The BE Journal of Economic Analysis and Policy, 16*(4), 12–23.
Meroni, E. C., & Vera-Toscano, E. (2017). The persistence of overeducation among recent graduates. *Labour Economics, 48*, 120–143.
O'Leary, N., & Sloane, P. (2016). Too many graduates? An application of the Gottschalk-Hansen model to young British graduates between 2001–2010. *Oxford Economic Papers, 68*(4), 945–967.
Oosterbeek, H., & Webbink, D. (2011). Does studying abroad induce a brain drain? *Economica, 78*, 347–366.
Ordine, P., & Rose, G. (2009). Overeducation and instructional quality: A theoretical model and some facts. *Journal of Human Capital, 3*(1), 73–105.
Ordine, P., & Rose, G. (2011). Inefficient self-selection into education and wage inequality. *Economics of Education Review, 30*(4), 582–597.
Ordine, P., & Rose, G. (2015). Educational mismatch and unemployment scarring. *International Journal of Manpower, 36*(5), 733–753.
Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. *American Economic Journal: Economic Policy, 3*, 148–171.
Oreopoulos, P., von Wachter, T., & Heisz, A. (2012). The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics, 4*(1), 1–29.
Parey, M., & Waldinger, F. (2010). Studying abroad and the effect on international labour market mobility: Evidence from the introduction of ERASMUS. *The Economic Journal, 121*, 194–222.
Puhani, P. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys, 14*(1), 53–68.
Redmond, P., & McGuinness, S. (2019). The gender wage gap in Europe: Job preferences, gender convergence and distributional effects. *Oxford Bulletin of Economics and Statistics, 81*(3), 564–587.
Reimer, D., Noelke, C., & Kucel, A. (2008). Labor market effects of field of study in comparative perspective: An analysis of 22 European countries. *International Journal of Comparative Sociology, 49*(4–5), 233–256.
Robst, J. (2007). Education, college major, and job match: Gender differences in reasons for mismatch. *Education Economics, 15*(2), 159–175.
Robst, J. (2008). Overeducation and college major: Expanding the definition of mismatch between schooling and jobs. *The Manchester School, 76*(4), 349–368.
Romanov, D., Tur-Sinai, A., & Eizman, G. (2017). Overeducation, job mobility and earnings mobility among holders of first degrees. *Applied Economics, 49*(26), 2563–2578.
Rubb, S. (2003). Overeducation: A short or long run phenomenon for individuals? *Economics of Education Review, 22*(4), 389–394.

Rubb, S. (2013). Overeducation, undereducation and asymmetric information in occupational mobility. *Applied Economics, 45*(6), 541–751.
Rubb, S. (2014). Factors influencing the likelihood of overeducation: A bivariate probit with sample selection framework. *Education Economics, 22*(2), 181–208.
Rumberger, R. W. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources, 22*(1), 24–50.
Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature, 31*, 831–880.
Sellami, S., Verhaest, D., Nonneman, W., & Van Trier, W. (2017). The impact of educational mismatches on wages: The influence of measurement error and unobserved heterogeneity. *The BE Journal of Economic Analysis and Policy, 17*(1), 20160055.
Sicherman, N. (1991). Overeducation in the labour market. *Journal of Labor Economics, 9*(2), 101–122.
Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy, 98*(1), 169–192.
Simon, C. J., & Warner, J. T. (1992). Matchmaker, Matchmaker: The effect of old boy networks on job match quality, earnings and tenure. *Journal of Labor Economics, 10*(3), 306–330.
Sloane, P. J., Battu, H., & Seaman, P. T. (1996). Overeducation and the formal education/experience and training trade-off. *Applied Economics Letters, 3*(8), 511–515.
Sloczynski, T., & Wooldridge, J. M. (2018). A general robustness result for estimating average treatment effects. *Econometric Theory, 34*(1), 112–133.
Thurow, L. (1975). *Generating inequality*. Basic Books.
Turmo-Garuz, J., Bartual-Figueras, M. T., & Sierra-Martinez, F. J. (2019). Factors associated with over-education among recent graduates during labour market integration: The case of Catalonia (Spain). *Social Indicators Research, 144*, 1273–1301.
Vera-Toscano, E., & Meroni, E. C. (2021). An age-period-cohort approach to the incidence and evolution of overeducation and skills mismatch. *Social Indicators Research, 153*(2), 711–740.
Verdugo, R. R., & Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources, 24*(4), 629–643.
Verhaest, D., Lavrijsen, J., Van Trier, W., Nicaise, I., & Omey, E. (2018). General education, vocational education and skill mismatches: Short-run versus long-run effects. *Oxford Economic Papers, 70*(4), 974–993.
Verhaest, D., & Omey, E. (2006). The impact of overeducation and its measurement. *Social Indicators Research, 77*, 419–448.
Verhaest, D., & Omey, E. (2010). The determinants of overeducation: Different measures, different outcomes? *International Journal of Manpower, 31*(6), 608–625.
Wiers-Jenssen, J., & Try, S. (2005). Labour market outcomes of higher education undertaken abroad. *Studies in Higher Education, 30*(6), 681–705.
Wooldridge, J. (2010). *Econometric analysis of cross-section and panel data* (2nd ed.). MIT Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.