DISCUSSION PAPER SERIES

IZA DP No. 7788

Overeducation at a Glance: Determinants and Wage Effects of the Educational Mismatch, Looking at the AlmaLaurea Data

Floro Ernesto Caroleo
Francesco Pastore

November 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Overeducation at a Glance: Determinants and Wage Effects of the Educational Mismatch, Looking at the AlmaLaurea Data

Floro Ernesto Caroleo
University of Naples "Parthenope"
and IZA

Francesco Pastore
Seconda Università di Napoli
and IZA

Discussion Paper No. 7788
November 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7788
November 2013

ABSTRACT

Overeducation at a Glance: Determinants and Wage Effects of the Educational Mismatch, Looking at the AlmaLaurea Data*

This paper provides the first available evidence on overeducation/overskilling based on AlmaLaurea data. We focus on jobs held 5 years after graduation by pre-reform graduates in 2005. Overeducation/overskilling are relatively high – at 11.4 and 8% – when compared to EU economies. *Ceteris paribus* they tend to be more frequent among children of parents with lower educational levels, through school tracking. Most arts degrees and social sciences, but also Geology and Biology are associated to both types of the educational mismatch. The quality of education is also a factor, suggesting that in addition to the low demand for skills, one should also reckon the inability of the educational system to provide work-related skills. Moreover, we find a non-conditional wage penalty of 20% and 16% and a conditional one of about 12% and of 7%, respectively. Heckit returns a sample selection corrected penalty slightly higher, supporting not only the job competition and job assignment models, but also the human capital model.

JEL Classification: C25, C26, C33, I2, J13, J24

Keywords: university-to-work transition, overeducation, overskilling, sample selection bias, AlmaLaurea, Italy

Corresponding author:

Francesco Pastore
Seconda Università di Napoli
Faculty of Law
Via Mazzocchi, 5
81055, Santa Maria Capua Vetere (CE)
Italy
E-mail: francesco.pastore@unina2.it

* Most part of this research was conceived during several periods of stay of Francesco Pastore at the AlmaLaurea headquarters in Bologna. We are grateful to Andrea Cammelli, Angelo Di Francia and Silvia Ghiselli not only for providing us with AlmaLaurea data, but also for detailed and very helpful suggestions on data analysis. Previous versions of this paper have been presented at: AlmaLaurea Conference, Bologna (2011), IAB of Nuremberg (2011), XXVI AIEL Conference, Catholic University of Milan (September 2011), University of Naples "Parthenope" (2011), University of Naples "Federico II" (2011), Second University of Naples (2011), University of Pescara (2012); University of Salerno (2013). We thank all seminar participants. We are also grateful to Lilia Costabile, Giuseppe Croce, Hans Dietrich, Francesco Ferrante, Claudio Lucifora, Nadia Netti, Patrizia Ordine, Marco Pecoraro, Claudia Pigini, Giuseppe Rose, Dario Sciulli, Francesca Sgobbi and Peter Sloane for valuable comments on earlier versions of this paper. This notwithstanding, the authors are solely responsible for the opinions expressed in this paper.

Introduction

It is difficult to say whether the supply of highly qualified labor is (and will be also in the future) greater, smaller or equal to its demand. In Italy, both the supply and demand of high skills are low by EU standards (OECD, various years). In spite of the low supply of graduates, though, overeducation and overskilling are rather widespread phenomena.

Overeducation is a cause of concern for the households and policy makers, because it causes a penalty to individuals in terms of earnings and employment opportunities and a waste of resources to the society as a whole in terms of underutilization of human capital and inefficiency of public expenditure on education (Groot 1996; McGuinness, 2006).

This paper provides the first available estimates of the impact of overeducation on wages of AlmaLaurea university graduates registered in the years before the implementation of the Bologna process back in 2001, so-called pre-reform graduates, and graduated in 2005. The focus of the empirical analysis is on their employment status 5 years after graduation.

We study the determinants of the mismatch by using the only questions available in the AlmaLaurea data. More specifically, overeducation is defined as referring to those individuals who declare that their university degree was neither required by law, nor useful to access their current job. Overskilling is defined as referring to those individuals who declare that the competences acquired during their university studies were not used in their current job.

The main point of this paper is that the relatively high share of overeducated/overskilled graduates could be due not only to an excess supply of graduates as compared to their low demand, which seems unlikely considering how small the supply is as compared to the European average, but rather to the lack of work related skills of young graduates and to the difficulty of the school to work transition system in spawning skills appreciated by the firms. In other words, we argue that the Italian case is perfectly in line with some recent interpretations of overeducation not as a breach of the human capital theory, but as indirect confirmation of its assumptions and conclusions (Leuven and Oosterbeek, 2011; Verhaest and van der Velden, 2013). We reach this conclusion by means of a new interpretation of the Heckman sample selection bias procedure as a test of the human capital explanation of overeducation.

The outline of the essay is as follows. Section one summarizes the relevant theoretical and empirical literature by focusing especially on the Italian case. Section two describes the AlmaLaurea data used in the empirical analysis and discusses the econometric methodology implemented. Section three analyses the determinants of the effectiveness of the university degree using ordered probit models, and the estimate of OLS conditional wage penalty controlling for the possible sample selection bias arising from measuring overeducation only

among the employed by the Heckit econometric specification of the earnings equation. The concluding section also discusses possible policy suggestions to reduce the impact of overeducation.

1. The state of the art

1.1 Theoretical explanations of overeducation

Theories that explain overeducation range between two opposite theoretical constructs: the human capital theory and the job competition model (Sloane 2003, McGuinnes 2006, Leuven and Oosterbeek 2011). Traditionally, overeducation has been considered an exception to the human capital theory as it is associated to a mismatch and therefore a market disequilibrium. Accordingly it should be a short term phenomenon as a sufficient degree of wage flexibility should restore any imbalance between supply and demand in the graduate labour market unless some persistent, often unobserved, low ability / skill problem affects the permanently overeducated. More recent literature tends to restore the validity of the human capital theory in explaining overeducation (Leuven and Oosterbeek 2011). As a matter of fact, it has been underlined that overeducation is rather a signal of a lack of the work-related component, rather than a waste of human capital. As it is well argued by the Nobel prize winner, Gary Becker (1964), human capital is not only represented by the level of education but also by generic work experience and that specifically acquired on a particular type of job. Overeducation is therefore a consequence of a lack of work experience and this is typical of young people, despite their increasing educational level.

The job competition model, brought to the fore, for the first time, by Lester C. Thurow (1975), on the other side, helps understanding the persistence of overeducation among the adults. In this case, excess schooling is a consequence of the competition for jobs in presence of rigidity of the demand for highly educated labour that leads graduates to accumulate education, which is in some cases more than that requested to get a job, in order to reach the best position in the queue for the job.

With the assignment theory, Sattinger (1993) attempted to reconcile the two previous theories. Like the job competition model, the model assumes that the jobs available in the economy are limited, which implies that remuneration is job specific and independent of the human capital endowment of the individual; on the other hand, like the human capital theory, it assumes that with their investment in human capital individuals are able to compete for the best

job and wages are bound to be influenced by the human capital level of individuals. Overeducation arises because wages will neither be entirely related to acquired schooling and other individual attributes, like in the human capital model, nor to the nature of the job, like in the job competition and job assignment model.

The job search theoretical model assumes, instead, that unemployment is largely a voluntary choice. People accept a job offer when it brings with it a wage higher than their reservation wages. The most skilled graduates prefer to wait into the non-employment pool until when they get the best job offer they can. High skill individuals have higher reservation wages and wait for a longer time than the least skilled graduates, who tend to choose the first job offer they get, even if it involves overeducation. Overeducation arises because the least skilled individuals get the first job offer they can because their reservation wage is low.

Overeducation may result also from career mobility theories: wages tend to grow over time together with the work experience accumulated by individuals. It is therefore physiological that firms and graduates generate job-worker matches with low earnings in the short run, but good career prospects in the long-run.

1.2. The empirical literature

While early studies focused on the USA (Freeman, 1976), more recently, overeducation and skill-mismatch patterns have been noted also in other economies, including several EU ones (see, for overviews, Büchel et al., 2003; Rubb, 2003; McGuinness, 2006; Quintini, 2011; Leuven and Osterbeek, 2011) and also Italy (see the references below). These studies have addressed, among others, the following broad issues:

a) Size and cross-country determinants of overeducation;
b) Individual characteristics of the overeducated/overskilled;
c) Penalty in terms of earnings and employment probabilities;
d) Shortcomings of OLS and corrections for measurement errors, sample selection and endogeneity bias.

Issue a) is one of the most complex to deal with, due to the lack of comparative data. A relevant problem arises from the scarcity of unambiguous definitions of over and/or under education, inviting to take the greatest caution especially in a comparative perspective (see, among others, Chevalier, 2003; Mavromaras et al. 2013; Sgobbi and Suleman, 2013; and the recent survey by Leuven and Oosterbeek, 2011).

Meta-analysis overviews of the empirical literature show that the share of overeducation among graduates ranges from 20 to 30% (Groot e van der Brink 2000; McGuinness, 2006; Leuven and Oosterbeek 2011; Croce 2012; Quintini 2011). Early evidence on the skill

mismatch across OECD countries (Manacorda and Petrongolo, 2000) point to higher overeducation in the EU as compared to the USA. This has been considered as a confirmation that technological change has been in the EU less skill biased than in the USA and has been developed in a low growth environment. This feature becomes more dramatic in Southern European countries, especially in Italy where there has been also a considerable increase of human capital supply (Cainarca and Sgobbi, 2009).

Comparative evidence based on REFLEX data reporting comparable information on graduates in the 1999/2000 academic year in 16 European countries (Davia et al., 2010; Verhaest and van der Velden, 2013) suggests that Italy has a higher than average share of overeducated workers. McGuinness and Sloane (2010, Table 3.6) find that, with a share of 23% of overeducated workers at the time of their first job and of 13% five years after graduation, Italy is the third last performer, standing only after Spain and the UK. In the other EU countries in the sample, overeducation is almost always under the 10% threshold. Slightly different is the case of overskilling, for which Italy tends to the EU average of 21% at the first job and 11% five years after graduation.

A recent, but flourishing stream of literature is attempting to estimate the relative impact of demand and supply side variables in cross-country panel data analyses. The authors find that demand side variables and differences in the imbalances between the composition by field of study of the demand for and supply of high education are important factors (Davia et al., 2010; Verhaest and van der Velden, 2013; Croce and Ghignoni, 2012; Ghignoni and Verashchagina, 2013). Interestingly, Verhaest and van der Velden (2013) find that the employment protection legislation is irrelevant, whereas the quality and orientation (general versus specific) of the educational system/program are important.

As to point b), overeducation is attributed to similar observed characteristics. In the case of Italy, Ferrante (2010, p. 89-93) uses AlmaLaurea data to assess the impact of a number of individual characteristics on the "effectiveness of the university degree"[1] in providing a job that is up to the educational and skill level of the individual. He reports that the variables that correlate positively and significantly with the above indicator are: a high school diploma with a grade of 55-60 out of 100; a high university final grade; a longer length of job search; experiencing some postgraduate training; holding a university degree in Engineering, Chemistry and Pharmacy, Law. The negative and statistically significant determinants include: holding a diploma of technical high school, rather than gymnasium; belonging to the working class; starting their career via starter or atypical working contracts, such as apprenticeship, stage, temporary contract; holding an arts degree or a degree in Education, Psychology, Social

[1] This is a special indicator elaborated by AlmaLaurea.

Sciences. Using the ISFOL-PLUS data, Franzini and Raitano (2012) add as individual factors that augment the probability to be overeducated: experiencing delayed graduation, having found their job through informal channels. Gender does not seem to affect the probability of overeducation. Cutillo and Di Pietro (2006) find that women have a lower probability of overeducation than men in the ISTAT survey of graduates. The university degrees most at risk are: Political Sciences, Literature and Languages. Instead, Law, Medicine, Sciences, Mathematics, Philosophy, Engineering, Architecture and Agriculture reduce the chance of overeducation with respect to Economics. Nonetheless, the ISFOL-PLUS data include many cohorts of age, making it hard to identify the impact of each degree in a given time, while ISTAT data have a smaller number of observations and greater heterogeneity of educational paths. Moreover, the AlmaLaurea data used here allow considering a number of factors which are bound to affect the risk of overeducation, such as pre- and post-graduate education and family background.

Even if the return to education (point c) is still positive for the overeducated as compared to secondary high school diploma holders (Wasmer et al, 2005; Brynin and Longhi, 2009; Franzini Raitano, 2012; Cainarca and Sgobbi, 2009), nonetheless, the overeducated invariably get a wage penalty as compared to their peers employed in positions for which they hold the required diploma (Sloane, 2003; Leuven and Oosterbeek 2011). Moreover, generally speaking, the wage penalty for overskilling is lower than that for overeducation (see, among others, McGuinness and Sloane 2010).

The wage penalty of overeducation / overskilling is found to be lower in Italy than in other countries and in some cases not statistically significant (Wasmer et al. 2005; Brynin and Longhi 2009; Ordine and Rose, 2009). Using the 2001 ISTAT enquiry on professional integration of 1998 graduates, Cutillo and Di Pietro (2006) find a wage penalty for university graduates ranging between 2.4% and 5.7%. McGuinness and Sloane (2010) find a wage penalty of about 10%. Interestingly, in the case of Italy, they find a higher wage penalty for the overskilled (-11%) than for the overeducated (-4%), the latter being not statistically significant. Using the ISFOL PLUS data, Aina and Pastore (2012) find a strong correlation of overeducation with delayed graduation and a wage penalty associated to overeducation of about 20%, slightly higher than in previous studies. The heterogeneity of current estimates suggests the need for reliable estimates, such as those based on AlmaLaurea data.

Many authors (point d) have raised the concern that simple OLS estimates tend to under/overestimate the wage penalty associated to overeducation. Three types of possible sources of bias have been discussed: a) measurement errors; b) endogeneity; c) sample selection. Measurement errors might tend to reduce the wage penalty since often individuals

believe, subjectively, more than they do objectively, to be overeducated (or also overskilled) when they are not. Chevalier (2003), Mavromaras et al. (2013) and Pecoraro (2011) elaborate ways to measure the wage effect of genuine versus apparent overeducation by looking at the relation between overeducation and job satisfaction. Unfortunately, the available data does not allow applying this definition.

Endogeneity arises if overeducation is assumed to be related to unobserved characteristics, such as a lower level of skills and motivation of the overeducated. Now, if the overeducated are less motivated than average, it is likely that the wage penalty is higher than that typically found. In fact, once controlling for unobserved motivation and skills, overeducation should generate a greater wage penalty.

While endogeneity tends to generate upward corrections of the wage penalty, and measurement errors tend to generate a downward correction, sample selection bias has a potentially ambiguous effect.

Nicaise (2001) is among the first to notice that ignoring the non-employed might generate a bias on returns to education whose direction is in principle ambiguous. Applying her line of reasoning to the case of overeducation, we may argue that according to the job competition, the job assignment and the human capital model, sample selection bias arises because of the fact that the educational mismatch appears first of all in the form of a higher probability of non-employment and only at a later stage it takes the form of a wage penalty. Once controlling for the selection bias arising from considering non-employment, the wage penalty of those experiencing the educational mismatch might be higher. Conversely, according to the search theoretical model, unemployment is a voluntary choice and the most skilled graduates prefer to remain non-employed waiting for the best job offer they can get. If employed, they would be less likely to experience overeducation. In this case, once controlling for the selection bias arising from considering non-employment, the wage penalty of those experiencing the educational mismatch might be lower.

Once controlling for endogeneity and sample selection bias, most authors find that the wage penalty associated to overeducation further increases lending support to the job competition, job assignment and human capital models. Using an ISTAT survey carried out in 2001 on graduates in 1998, Cutillo and Di Pietro (2006) find that the wage penalty associated to overeducation increases up to between 22 and 39%, once controlling for endogeneity, and up to about 40%, once controlling for both endogeneity and sample selection bias.

2. Data and variable definition

This paper provides the first available estimates of the wage effect of overeducation on wages of AlmaLaurea university graduates. AlmaLaurea is a consortium including a large and growing number of Italian universities[2]. The aim of the consortium is to provide a framework to ease the interaction of graduates and firms by collecting the curricula of graduates and making them available to firms wishing to fill in their job vacancies. A further support is supplied to the universities providing them with homogenous information on the quality of the achieved education. It also collects valuable information on individual and educational characteristics of graduates at the time of graduation and on their employment status after one, three and five years from graduation.

By collecting very detailed information on several aspects of university education and the school-to-work transition of graduates, for each of the universities joining the Consortium, AlmaLaurea is the most important source of information to assess the quality of tertiary education in a comparative perspective across athenaeums, faculties, provinces, fields of studies and so on. The quality of education can be assessed by looking not only at course attendance and other parameters regarding the university performance of graduates, but also at their early labor market performance.

The sample on which this study is based includes all pre-reform university graduates who obtained their degree in 2005 at one of the 36 universities belonging to the Consortium at that time. As such, our sample includes quite a large number of *fuoricorso*, namely individuals who get their degree some years after the prescribed period. This is likely to generate some overestimation of the share those experiencing the educational mismatch and the associated wage penalty. Nonetheless, *fuoricorsismo* is a common phenomenon in Italy, regarding about 40% of graduates (Aina et al., 2013), and therefore focusing on the post-reform graduates would have excluded the *fuoricorso*, causing even a greater bias[3].

Individuals in the sample are observed at the time of their graduation and, thereafter, in 2006, in 2008 and in 2010. Our focus is on individuals who answer the questionnaire 5 years after graduation, which should allow us catching the determinants of permanent, rather than

[2] In 2013, 64 universities were members. All Italian regions are covered, except for Lombardy, whose universities have established their own consortium, called Stella. For further details about the Consortium, see the homepage: http://www.almalaurea.it/info/chisiamo.

[3] The 3+2 reform was implemented in 2001. Considering that most graduates obtain their degree in more than 7-8 years, it is likely that a more representative sample would be available only in the last years of the 2000s, which is too late for observing them five years from their degree.

transient overeducation. Our data should not be affected by the Great Depression which produced its main effect on the Italian labor market only starting from the end of 2011.

The sample consists of 28,976 pre-reform graduates interviewed at the time of graduation, 21,605 of whom answer the questionnaire 5 years after graduation and 17,387 of whom report being employed. The earnings equations include only 16,591 graduates because 796 employees do not report their wage, either for boredom or because it is too low to report.

The employment questionnaire administered after graduation includes two questions that provide subjective measures of the educational / skill mismatch in as much as they are based on individual' self-assessment. Question A16 of the questionnaire asks: "In your current job, do you use the competences acquired during your university studies?" Three answers are possible: 1) the competencies acquired are used to a great extent; 2) they are little used; 3) they are not used at all. We defined as overskilled those who choose answer 3. This question closely mirrors what Dolton and Silles (2008) call the "to do" definition of the mismatch.

Question A17 asks: "Is your university degree necessary to access your current job?" Four answers are possible: 1) the degree is required by law; 2) it is not required by law, but is in fact needed; 3) it is not required by law, but is in fact useful; 4) it is neither required by law nor useful. We defined as overeducated all those who choose answer 4. This question allows us defining what Dolton and Silles (2008) call the "to get" definition of the mismatch.

Earnings are defined as the natural logarithm of net monthly wages. Question A20 asks the interviewee to declare to which of 13 classes of €250 of monthly earnings, up to the "over €3,000" class, (s)he belongs to. For ease of analysis, the natural logarithm is applied to the average value of the relative class. In our earnings estimates, we use both the OLS and interval regression method. No information on working hours is available.

Such individual characteristics as civil status are asked only at the time of graduation, not five years later. This secures the exogeneity of these variables, but cannot prevent them from being inaccurate in the case of civil status, for instance.

The independent variables are self-explaining. They have been grouped in: a) individual characteristics (gender, civil status, having children, nationality); b) educational background (type of high school); c) university attendance and performance (final grade, time to get a degree, field of study); d) pre-graduation work experience (experience of study abroad, work and study); e) post-graduate studies and training; f) whether the graduate moved from the place where she got her degree. Table A1 in the Appendix provides the descriptive statistics for the overeducated/overskilled and the rest of the sample.

3. Results

3.1. Size and composition of the sample

The AlmaLaurea databank confirms a picture where the share of the overeducated/overskilled is roughly similar to that found in previous studies, but with a level of detail and accuracy never provided before. Table 1 shows that one year after graduation, the overskilled and the overeducated amounted to about 16.5 and 13.2 per cent respectively, to reduce at a roughly constant pace down to 11.4 and 8.0 per cent respectively at the end of the considered period. It means a reduction down only to 69.1% and 60.6% of the original value. In other words, for a large number of individuals it is *not* a transitory phenomenon. That seems to confirm only partially the career mobility theory. Interestingly, the AlmaLaurea figures are close to those of McGuinness and Sloane (2010) reported in a previous section.

[Table 1 about here]

The descriptive evidence highlights a massive and quite generalized disruption of the human capital that the university system has generated. Figure 1 reports the non-employment shares by field of study, the share of overeducation and that of the well-matched individuals 5 years after getting the degree. The non-employment share is dramatically high for most degrees, with an average of 40%. It ranges between a minimum of 27% in the case of Engineering and a maximum of about 50% in the case of Geology and Biology. The low employment share of graduates in Medicine is due to the fact that most medical doctors are still involved in postgraduate schools.

A more accurate measure of the human capital waste is given by the sum of non-employment and overeducation/overskilling. Interestingly, it is quite common that the higher is non-employment share, the higher is also the share of the overeducated.

Figure 2 displays the share of overeducated and overskilled for each type of degree. The field of study is undoubtedly the main factor. Overeducation ranges between zero or nearly zero in the case of Medicine, Architecture, Chemistry and Pharmacy, Engineering and Sciences and more than 10% in the case of Geology and Biology (10.2%), Physical Education (12.2%), Languages (13.2%), Political and social sciences (14%) e Literature (17.9%). Overskilling follows about the same pattern with a share slightly higher for every type of degree. Languages (16.5%), Political Sciences (18.4%), Geology and Biology (18.7%), Physical Education (20.7%) and Literature (25%) present the more remarkable share even if it is difficult to say if this is overall high or low.

Table A1 in the Annex contains a more detailed picture of the productivity characteristics of the overeducated and overskilled versus the perfectly matched. The table suggests that the former have much lower human capital level than the latter, which we explore in more depth in the next section.

[Figure 1 and 2 about here]

3.2. Determinants

Table 2 reports the results of logit estimates of the probability to be overeducated (column 1) or overskilled (column 2). The table reports odds ratios of the independent variables[4], rather than estimated coefficients. Overall, the estimates are quite satisfactory for these types of cross-section data, with values of the pseudo-R2 of about 0.10. The share of correctly classified cases is satisfactory and so is also the area under ROC curve.

Gender is a statistically significant (at a level of 5%) determinant of overskilling, even in conditional estimates controlling for a number of variables, but not of overeducation. Women are about 13 odds points more likely to be overskilled. In fact, unreported estimates confirm that women have a statistically significant and higher non-conditional probability of both overeducation and overskilling, by about 15 and 36 odds points, respectively. The gap dramatically reduces, disappearing in the case of overeducation, when we include controls for human capital variables, suggesting that the highest non conditional probability of mismatch of women is due to their tendency to concentrate in those fields of study which are more subject to the mismatch.

Other individual characteristics, such as the civil status and having children at the time of graduation seem to have little impact on the probability to be overeducated. A possible explanation is that, as already noted, the civil status might not apply anymore to the individual 5 years after graduation.

The statistical significance of variables catching secondary high school performance and of the school tracking coefficients seem to confirm what a large literature says about the role of the family socio-economic background in affecting the educational choice and the performance of

[4] The odds ratio is the ratio of the odds of an event occurring in one group to the odds of it occurring in another group, namely the default category. If the probabilities of the event in each of the groups are p_1 (first group) and p_2 (second group), then the odds ratio is:

$$\frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{p_1/q_1}{p_2/q_2} = \frac{p_1 q_2}{p_2 q_1},$$

where $q_x = 1 - p_x$. An odds ratio of 1 indicates that the condition or event under study is equally likely to occur in both groups. An odds ratio greater than 1 indicates that the condition or event is more likely to occur in the first group. And an odds ratio less than 1 indicates that the condition or event is less likely to occur in the first group.

graduates on the labor market. The same groups that are at disadvantage in achieving higher education are at disadvantage also in their access to the labor market and, we add here, tend to experience very often also overeducation and / or overskilling. The common factors of success can be found in the socio-educational background, which affects educational choice and labor market outcomes through school tracking, as now a large literature has proven (see, among others, Brunello and Checchi, 2007; Caroleo and Pastore, 2012). In fact, although being totally free, the choice of the type of high school tends to reflect the social class to which the young person belongs to: people from a poor walk of life tend to choose technical or professional schools and, hence, to experience problems in their educational career later on, augmenting also the chance of being overeducated in their working life. For these reasons, once controlling for their performance at high school, it should come as no surprise that the educational background of parents does not affect directly the probability to be overeducated/overskilled. We have hence dropped family educational and social background from our final estimate.

Confirming a finding relative to the USA, UK and Germany (for a survey of recent contributions, see Leuven and Oosterbek, 2011), several aspects of an individual's educational quality correlate with the likelihood to experience overeducation / overskilling. This is the case of the field of study, the final grade, the time spent to get a degree. Confirming a finding of Aina and Pastore (2012) based on the ISFOL data, *fuoricorso* graduates with two or more years of delay have ceteris paribus a 50 odds point greater chance than the *in corso* to experience the educational mismatch.

The impact of the field of study is particularly important, also ceteris paribus. All fields of study are associated with a higher chance of overeducation / overskilling than engineering (the reference group). Particularly strong is, ceteris paribus, the impact of holding a degree in Literature, Languages, Physical Education, Political and Social Sciences, Psychology and Geology and Biology. Architecture and Medicine present coefficients statistically not different from Engineering. Overall, the quality of education, as measured by indicators of university performance, seems to be the most important determinant of the probability of experiencing overeducation/overskilling. We will come back on this crucial point when discussing the results.

The localization of job search matters. The graduates who seek their job in the North, no matter whether in the West or East, experience a much lower probability of overeducation/overskilling than their pears searching in the Centre and even more so in the South. Moving abroad reduces the risk of overeducation, but not of overskilling, probably because of the country-specific content of some of the skills acquired in the educational system, which makes them less easy to transfer.

Furthermore, our estimates deliver useful information to the policy maker, while providing support to our interpretation of the educational mismatch. Interestingly, on-the-job training practices, attending some graduate schools, II level master degrees, namely those that can be accessed after getting the full specialist degree – either the traditional 4-year program, the 3+2 or the 5-year program (so-called *laurea specialistica a ciclo unico*) – reduce the risk of overeducation/overskilling in a statistically significant manner. I level master programs reduce the risk of overeducation, but not overskilling, which are positively affected by post-degree scholarships. Other post-degree programs – such as the doctorate, other types of master degrees, stages, public training programs and the voluntary social service – are not statistically significant.

On the whole, estimates provide useful information to policy makers at all levels of the educational system about the fields of study with a lesser risk of mismatch. As to the interpretation, our findings seem to indirectly confirm, rather than being in contrast with the human capital model. In fact, overeducation/overskilling are strongly associated with characteristics which denote a low quality of human capital. In other words, while increasing in the educational dimension, the human capital of Italian graduates does not sufficiently develop in terms of the competences which are directly requested in the labor market. In turn, this tends to affect the ability of graduates to access high quality jobs. In other words, there is a demand for job specific competences in the labor market which remains unsatisfied.

[Table 2 about here]

3.3. The wage effect

A simplified variety of the Verdugo and Verdugo (1989, cit. in Leuven and Oosterbeek, 2011) version of the classical ORU specification (Over-, Required, and Undereducation), in which the usual Mincerian earnings equation is augmented by overeducation and/or overskilling dummies, provides the empirical framework to estimate the wage penalty[5]. In other words, the following equation has been estimated

$$Lnw_i = r^{OLS} O_i + \sum_{i,j=1}^{n} \beta_j X_{i,j} + u_i \qquad [1]$$

where $Ln\ w_i$ is the natural logarithm of the net monthly wage for an individual i, the X_i are a set of control variables assumed to affect earnings (see the previous sub-section for a full list of

[5] Unfortunately, the years of overeducation / overskilling are not available in the AlmaLaurea data, although they should be preferred according to Leuven and Oosterbeek (2011, section 5). As it is clear from the definition reported in the data section, the data do not allow measuring under-education either.

regressors) and the β_j are their coefficients. O_i is a dummy taking a value of one when the individual i is overeducated / overskilled and r^{OLS} is the estimated coefficient. u_i is a disturbance term representing other forces which may not be explicitly measured, assumed independent of X_i and O_i. To reduce endogeneity problems to a minimum and simplify the comparison of findings with the model with sample selection bias correction, we exclude the job characteristics.

Table 3 provides summary measures of the wage gap derived from the different estimated models, including the unconditional estimate, the conditional one (based on Table 4) and that obtained including controls for sample selection bias (based on Table 6). The unconditional wage penalty has been estimated both using a traditional OLS specification and the regression with intervals considering how wages are measured in the data.

The unconditional wage gap is relatively high for both overeducation (from -21 to -25%) and overskilling (from -16 to -21%). In both cases, OLS underestimates the wage penalty as compared to interval regression.

[Table 3 about here]

However, the unconditional measure of the wage penalty might catch such factors as the lower than average productivity characteristics of the overeducated / overskilled. In other words, such a high unconditional penalty might disappear once controlling for the lower than average levels and quality of human capital of the overeducated. Such characteristics might be observed or unobserved.

Table 3 reports also conditional measures of the wage penalty as obtained in OLS estimates and in interval regressions including all the variables of the AlmaLaurea data base as controls (as in Table 4). Interestingly, once controlling for the level and quality of human capital, both OLS coefficients are halved. More precisely, the wage penalty of overeducation reduces to 12% and that of overskilling to 7%. Similar reductions are observed in the case of interval regressions.

How to explain this result? As already shown in the previous section, the observed quantity and quality of human capital that the overeducated and the overskilled possess are actually lower than average. This partly explains their lower earnings. Table 4 presents detailed OLS estimates of the earnings equations augmented of the overeducation and overskilling terms[6].

[Table 4 about here]

As a further check, we test whether the impact of overeducation on wages is independent of that of overskilling. In other words, which of them matters more in terms of wage penalty? Table 5 attempts to answer this question by reporting coefficients of OLS unconditional and

[6] The interval regressions, which are similar to the OLS, are available from the authors on request.

conditional earnings equations, where there are three variables for the educational mismatch: overeducation only, overskilling only and overeducation and overskilling. The results suggest that the weakest group is represented by those who are at the meantime overeducated and overskilled. They have an unconditional wage penalty of about 24% and a conditional wage penalty of half that size. This result holds also in interval regressions. Overskilling only is almost entirely caught by observed characteristics of individuals, such as the performance at school and at the university in OLS estimates, while it still bears a 5% wage penalty in interval regressions.

[Table 5 about here]

3.4. The Heckit estimates

As already noted, OLS estimates do not control for possible unobserved differences between the overeducated and the non-employed, who might also experience the educational mismatch if employed. We control for the possible omitted heterogeneity bias arising from measuring overeducation only among the employed by the Heckman (1979) econometric specification – sometimes called Heckit for consonance to the Tobit model – of the earnings equation, where the usual OLS estimates are corrected for the lower/higher employment opportunities of the most skilled and motivated among those whose personal attributes would lead to overeducation if they were employed.

In analytical terms, and in short, equation [1] would be not correctly specified, since it does not consider the existence of another variable, which, if statistically significant, might return a biased estimate of the coefficient of interest, *r*:

$$Lnw_i = r^{Heckit} O_i + \sum_{i,j=1}^{n} \beta_j X_{i,j} + \rho\lambda\left(\sum_{l=1}^{m} \theta_l Z_{i,l}\right) + u_i \qquad [2]$$

where *r* is now denoted with the superscript Heckit, to distinguish it from the corresponding OLS estimate; ρ is the correlation between the error terms of the main and of the participation equation and λ is the inverse Mills ratio evaluated at the mean of the covariates ($Z\vartheta$), which include in addition to the *X*, also one or more instrumental variables. When there is sample selection bias, the latter term should be included in the earnings equation to obtain unbiased estimates of the parameters of interest.

Now, two possibilities are in order:

H_0: $r^{Heckit} > r^{OLS}$ [3]

H_1: $r^{Heckit} \leq r^{OLS}$

As shown in panel (a) of Figure 3, according to H_0, OLS is underestimating the wage penalty associated to overeducation/overskilling. Only the most skilled overeducated (pink dots) are selected into employment. Once controlling for the least skilled and motivated among those experiencing the educational mismatch, the coefficient of the overeducation variable increases (green line). As noted in the survey section, H_0 is consistent with the job competition and the job assignment model, whereas unemployment is high and hence dominated by the involuntary component. The most skilled are the first to get job offers and accept them as the best alternative. We argue that it is consistent also with the human capital model: the least skilled would be so because of their lack of work related skills.

As shown in panel (b) of Figure 3, according to H_1, OLS is overestimating the wage penalty associated to overeducation/overskilling. Only the least skilled overeducated are selected into employment (pink dots). Once controlling for the most skilled and motivated among those experiencing the educational mismatch, the coefficient of the overeducation variable shrinks (green line). H_1 is consistent with the search theoretical models, whereas unemployment is assumed to be voluntary in nature and the most skilled graduates prefer to wait in the unemployment pool for the best job offer to come.

[Figure 3 about here]

Table 6 reports the results of earnings equations estimated with the Heckman correction. A maximum likelihood simultaneous estimate is preferred to the two step procedure. The main equation is a typical Mincerian earnings equation, while the selection equation is a probit estimate of the probability to be employed rather than non-employed.

We apply the rule by which variables in the main and selection equation should be the same, except for some instrumental variables (Cameron and Trivedi, 2009; Wooldridge, 2003). The instrumental variables have been chosen according to the criterion that they should affect the probability to participate to the labor market, but not wages. This implies that the set of regressors used here are only a sub-set of those used in the OLS equation.

The instrumental variables typically used to predict labor force participation in these cases, namely civil status and having children, in our estimates of Table 4, influence wages as well, although they are not statistically significant determinants of overeducation and overskilling. That does not allow us using these variables as instruments in the selection equation. We have, hence, included these variables in both selection and main equation. As alternative instruments, we use the educational level of parents, based on the assumption that differences among graduates in the probability of finding a job are essentially linked to the socio-economic background of parents (see, among others, Brunello and Checchi, 2007; Bratti, Checchi and De Blasio, 2008; Caroleo and Pastore, 2012). This is a consequence of the weakness of other tools

of job search that are able to equalize chances in the labor market. In the absence of other plausible instruments, we have also exploited as instrument the strong non linearity of the ML function, as suggested by Cameron and Trivedi (2009).

Table 6 (and Table 3) give the results of the Heckit estimates using as instruments the educational level of fathers and mothers. The results relative to the estimate with different definitions of overeducation/overskilling are given in the last columns of Table 5. The arthrho variable, which indicates the correlation between the two equations, is negative and statistically significant, although to a lesser extent in the case of overskilling. The Wald test of independence between the main and selection equation confirms this result: in fact, it rejects the hypothesis H_0 with a high level of significance in the case of overeducation, but only at a 10% significance level in the case of overskilling. Overall, this suggests that there are unobserved factors that affect the labor participation chances and, consequently, the reservation wage and also the wage received if employed. The same result is obtained with the two step estimates.

[Table 6 about here]

Nevertheless, similar to Cutillo and Di Pietro (2006), the wage penalty associated to overeducation/overskilling augments by only about 1% in both cases. This finding could be partly due to the inadequacy of the available instruments in fully correcting for sample selection bias or also to the extreme flatness of the (entry) wage distribution especially among a homogeneous sample like the one considered here. Taken at its face value, though, overall, our findings can be seen as weak evidence in support to the job competition and the job assignment models instead of the job search model. In fact, it is also in line with our interpretation of the Italian overeducation as based on the human capital model. In other words, the non-employed if employed would have a higher chance of overeducation/overskilling also due to their low human capital level, especially the work-related components.

4. Discussion, concluding remarks and policy implications

This paper has attempted to study the main determinants and labor market effects of overeducation/overskilling using, for the first time, the AlmaLaurea database. This is by far the largest databank of graduates in Italy. The sample includes the universe of pre-reform graduates from the 36 universities belonging to the consortium in 2005. The focus is on labor market outcomes five years from graduation. The data allows us establishing a number of links with overeducation that have never been investigated before with the same detail. Results show that overeducation and overskilling are persistent phenomena.

We argue that the evidence provided in this paper is consistent with a new interpretation of the educational mismatch in Italy. The usual interpretation is that they are due to a lack of demand for human capital for the country is still strongly specialized in the traditional manufacturing industry, but is experiencing a dramatic increase in the share of graduates. This view is questioned in our analysis. We think that the inefficiencies of the tertiary educational and training system and, particularly, the difficulties in enhancing job related competences of youth also matter. This interpretation is consistent with the recent theories of the educational mismatch that trace it back to the too low human capital, rather than to an excess of it, since, in spite of the growing level of youth education, job competences continue to be insufficient (Leuven and Oosterbeek, 2011). Our interpretation does not neglect the importance of the demand side. Our analysis assumes that there is a potential demand for skills in the production system which remains unexploited because of the youth experience gap and their educational mismatch.

This interpretation is based, above all, on evidence coming from analysis of the determinants of overeducation. The factors that are associated to overeducation are consistent with the well-known image of an immobile social structure, whereas not only success at school and at the university, but also in the labor market dramatically depends on the socio-educational background of young people, which strongly affects the choice of the field of study, as a large literature has ascertained (see, among others, Checchi et al., 1999; Caroleo and Pastore, 2012). The chances of overeducation/overskilling are strongly associated with any other university degree but Engineering, Medicine and few others. Particularly strong is the impact of holding a degree in Social Sciences, but also in some scientific fields such as Geology and Biology.

On the other hand, having completed some post-graduate training or advanced master courses, especially those involving on-the-job training, represents a cushion against the risk of overeducation, confirming the lack of job specific competences of graduates.

The correlation between overeducation, but even more so overskilling, on the one hand, and weak educational background and poor university performance, on the other hand, suggests, in fact, that overskilling is a signal of low skills. In other words, even when there is the right match between the qualification held by the graduate and that required to get the job, perhaps due to scant work experience, some graduates might be given tasks for which they feel to be overskilled.

The unconditional wage penalty associated to the educational mismatch is higher than that found in other similar studies, relatively more for overeducation (between -21 and -25%) than for overskilling (between -16 and -21%). Correcting for the observable characteristics available in the AlmaLaurea data base in a multivariate context, the wage gap associated to both forms of

educational mismatch halved. This mirrors the low human capital endowment of the overeducated/overskilled as compared to the individuals with the required level of skills and this explains also the most important part of the wage gap.

The Heckit correction has been used to control for the possible sample selection bias arising from measuring overeducation only among the employed without considering the different characteristics that modify the probability to be overeducated among the non-employed. The Heckit correction confirms that there is negative selection into employment of the most skilled among those experiencing the educational mismatch, as in the job competition and job assignment models. We argue that this conclusion is also consistent with the human capital model. Nonetheless, in our data, the wage penalty augments only of about 1%. This may be due to the high youth unemployment rate, which weakens the selection mechanism or also to the smoothness of the entry wage distribution.

Overall, the findings of this paper have important policy implications. From the demand side, they suggest that the most important strategy to reduce the share of overeducation and the wage penalty associated to it would be that the country move away from a low towards a high road to development. This is overall the most important strategy to accommodate the increasing supply of human capital of the youngest generation. This would also reduce the share of graduates who possess attributes that are not on demand in the labor market and experience a greater difficulty in finding any kind of job.

Related to this point, as, among others, Ferrante (2010) note, it is important to mention the small average size of Italian firms. Small firms do not manage human resources in such a way to fully exploit and develop them. This is due to their typically informal production structure, the scant propensity to delegate functions to managers, and the lack of on-the-job training programs.

From the supply side, it is important to: a) increase the quality of tertiary education and of human capital in general; b) reduce the length of studies for the individuals coming from low family background, so to reduce the impact of school tracking on university success; c) provide more guidance for families and students when deciding their field of study at the university; d) provide vocational education and training also at a university level for individuals with low family background (German solution): this implies adopting the dual principle on a large scale and providing on-the-job training before or soon after the university degree; e) fully implement the Bologna process.

References

Aina, C. and F. Pastore (2012), "Delayed Graduation and Overeducation: A Test of the Human Capital Model versus the Screening Hypothesis", *IZA discussion paper*, n. 6413, March.

Aina, C., Baici, E., Casalone, G. e Pastore, F. (2013), «Il fuoricorsismo fra falsi miti e realtà», Economi Lavoro, XLVII(1), 147-154.

AlmaLaurea (2005), Lavorare dopo la laurea. Caratteristiche e percorsi occupazionali, Il Mulino, Bologna.

Becker, G.S. (1964), *Human Capital*, New York, Columbia University Press.

Bratti M., Checchi D. and G. de Blasio (2008), "Does the Expansion of Higher Education Increase the Equality of Educational Opportunities? Evidence from Italy", *Labour*, 22(special issue): 53-88.

Brunello, G. And D. Checchi (2007), "Does School Tracking Affect Equality of Opportunity? New International Evidence", *Economic Policy*, 22(52): 781-861.

Brynin M., Longhi S. (2009), "Overqualification: Major or minor mismatch?", Economics of Education Review, 28(1): 114-121.

Büchel, F., A. de Grip and A. Mertens (eds), Overeducation in Europe, Edward Elgar, Cheltenham.

Cainarca G.C. and F.Sgobbi (2009), "Quanto paga studiare: la relazione fra istruzione e retribuzione in Italia", Economia Politica, XXVI(1): 53-84.

Cameron, A.C. e Trivedi, P.K. (2009), *Microeconometrics Using Stata*, College Station, TX, Stata Press, New York.

Caroleo, F.E. and F. Pastore (2012), "Talking about the Pigou Paradox. Socio-Educational Background and Educational Outcomes of AlmaLaurea", International Journal of Manpower, 33(1): 27-50.

Chevalier, A. (2003), "Measuring Over-education", Economica, 70(279): 509–531.

Croce, G. (2012), «L'overeducation in Europa e in Italia. Dobbiamo prenderla sul serio?», Meridiana, 70-72, 75-100.

Croce, G. and E. Ghignoni (2012), "Demand and Supply of Skilled Labour and Overeducation in Europe: A Country Level Analysis", Comparative Economic Studies, 54(2): 413-439.

Cutillo, A. and G. Di Pietro (2006), "The effects of overeducation on wages in Italy: a bivariate selectivity approach", International Journal of Manpower, 27(2): 143-168.

Davia M.A., McGuinness S., O'Connell P.J., 2010, "Explaining international differences in rate of overeducation in Europe", ESRI, working paper, n° 365,.

Di Pietro, G. and P. Urwin (2006), "Education and Skills Mismatch in the Italian Graduate Labour Market", Applied Economics, 38(1): 79-93.

Dolton, P. and M.A. Silles (2008), "The Effects of Overeducation on Earnings in the Graduate Labour Market", Economics of Education Review, 27(2): 125-139.

Ferrante, F. (2010), "«Educational / skill mismatch»: cosa ci dicono i dati AlmaLaurea?", in AlmaLaurea, XII Rapporto sulla condizione occupazionale dei laureati. Investimenti in capitale umano nel future di Italia ed Europa, Il Mulino, Bologna, 101-115.

Franzini M. and M. Raitano (2012), Few and underutilized? Overeducation of Italian graduates, in Mandrone E. (ed) Labour Economics: PLUS Empirical Studies, ISFOL, Temi e Ricerche 3, Ediguida, Cava dè Tirreni, ITA.

Freeman, R. (1976), The Overeducated American, Academic Press, New York.

Ghignoni, E. and A. Verashchagina (2013) "Educational qualifications mismatch in Europe. Is it demand or supply driven?", forthcoming in *Journal of Comparative Economics*.

Groot, W. (1996), "The incidence of, and returns to overeducation in the UK", Applied Economics, 28(10): 1345-1350.

Groot, W. and Maassen van den Brink, H. (2000), "Overeducation in the labour market: a meta-analysis", Economics of Education Review, 19(2): 149-58.

Heckman, J.J. (1979), «Sample selection bias as a specification error», *Econometrica*, 47(1), 153-62.

Leuven E. and H. Oosterbeek (2011), "Overeducation and Mismatch in the Labour Market", in E.A. Hanushek, S. Machin and L. Woessmann, Handbook of the Economics of Education, vol. 4: 283-326.

Manacorda, M. and B. Petrongolo (2000), "Skill Mismatch and Unemployment in OECD Countries", Wirtschaftspolitische Blatter, 47(1): 72-82.

Mavromaras, K. G., McGuinness, S., O'Leary, N. C., Sloane, P. J., and Wei, Z. (2013), "Job mismatches and labour market outcomes: Panel evidence on university graduates", Economic Record, 89(286): 382-395.

McGuinness, S. (2006), "Overeducation in the labour market", Journal of Economic Surveys, 20(3): 387-418.

McGuinness, S. and P.J. Sloane (2010), "«Educational / skill mismatch» nel mercato del lavoro dei laureate: un'analisi comparata", in AlmaLaurea, XII Rapporto sulla condizione occupazionale dei laureati. Investimenti in capitale umano nel future di Italia ed Europa, Il Mulino, Bologna, 101-115.

Nicaise, I. (2001), "Human capital, reservation wages and job competition: Heckman's lambda reinterpreted", Applied Economics, Vol. 33 No. 3, pp. 309-15.

OECD (various years), *Education at a Glance: OECD indicators. Country note: Italy*, Paris.

Ordine, P. and G. Rose (2009), "Overeducation and Instructional Quality: a Theoretical Model and Some Facts", Journal of Human Capital, 3(1): 73-105.

Ortiz, L. (2010), "Not the right job, but a secure one: overeducation and temporary employment in France, Italy and Spain", Work, employment and society, 24(1): 47-64.

Pecoraro, M. (2011), "Is there Still a Wage Penalty for being Overeducated but well-matched in Skills? A Panel Data Analysis of a Swiss Graduate Cohort", ," Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2011019, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Quintini, G. (2011), *Over-Qualified or Under-Skilled: A Review of Existing Literature*, OECD Social, Employment and Migration Working Papers, 121, Paris.

Rubb, S. (2003), "Overeducation in the Labour Market: A Comment and a Re-Analysis of a Meta-Analysis", Economics of Education review, 22(6): 621-629.

Sattinger, M. (1993), "Assignment models of the distribution of earnings", *Journal of Economic Literature*, 31(1): 851-80.

Sgobbi F., Suleman F. (2013), "A methodological contribution to measuring skill (mis)match", *The Manchester School*, 81(3): 420-437.

Sloane, P. (2003), "Much Ado about Nothing? What does the Overeducation Literature really Tells us?", in F.A. Büchel, F., A. de Grip and A. Mertens (op. cit.).

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

Verhaest D., van der Velden R. (2013), "Cross-country differences in graduate overeducation", *European Sociological Review*, 29(3): 642-653.

Wasmer E., Fredriksson P., Lamo A., Messina J., Peri G. (2005), The Macroeconomics of Education, Fondazione Rodolfo De Benedetti Report, Milan.

Wooldridge, J. (2003). *Introductory Econometrics: A Modern Approach*. Thomson Learning: Ohio.

Appendix of Tables and Figures

Table 1. Overeducation after 1, 3, 5 years of pre-reform graduates

Definition	1 year	3 year	5 year
Overskilled ("to do" definition)	16.47	12.49	11.44
Overeducated ("to get" definition)	13.16	9.37	7.99
Number of observations	13500	17223	17387

Source: own elaboration on AlmaLaurea data.

Table 2. Determinants of overeducation 5 years after graduation. Pure pre-reform graduate in 2005. Log odds ratios

	Logit	Logit
Dependent variable	Overeducation ("to get")	Overskilling ("to do")
Independent variables	(1)	(2)
Individual Characteristics		
Gender. Default: Men		
Woman	0.9112	1.1277**
Civil Status, Male. Default: Single		
Married	1.3706	1.1783
Lives together	1.1641	0.9821
Separated, divorced, widow	0.564	0.9668
Civil Status, Female. Default: Single		
Married	0.9802	0.9523
Lives together	1.0728	0.92
Separated, divorced, widow	0.727	0.3719**
Number of children, Men	0.5841*	0.7369
Number of children, Women	0.8543	1.051
Non Italian	1.0766	1.0438
High secondary School		
Type of Secondary high school diploma. Default: Lyceum in sciences		
Classical high school	0.9079	1.0213
Specialisation in teacher training	0.7900*	1.003
Language high school	1.2582*	1.0705
Art school	1.7490***	1.3975*
Technical school	1.3541***	1.1776**
Professional school	1.5591***	1.6706***
Other high school diploma	1.4884*	1.2664
University performance		
Final grade at the university. Default: Magna cum laude		
66-90 out of 110	1.9540***	2.2542***
91-100 out of 110	1.8596***	1.9335***
101-105 out of 110	1.5401***	1.8858***
106-110 out of 110	1.4571***	1.5971***
Time to get a degree. Default: Curricular years		
I extra-curricular year late	1.3314*	1.3141*
II extra-curricular year late	1.5746***	1.3815**
III extra-curricular year late	1.5557**	1.6536***
IV extra-curricular year late	1.5520**	1.5552***
V extra-curricular year late	1.9378***	1.8258***
Field of study. Default: Engineering		
Agriculture	4.4684***	3.7624***
Architecture	1.339	1.3471
Economics and statistics	3.4690***	1.6251***
Physical education	9.0089***	7.8656***
Geology and biology	7.0864***	6.2336***
Law	3.6917***	3.2842***
Education	4.3579***	2.4162***
Arts	15.3943***	9.8837***
Languages	9.1920***	5.2738***
Political and social sciences	9.5990***	5.5167***
Psychology	9.8709***	5.9034***
Mathematics and Physics	2.5189***	2.8105***

Study abroad. Default: No study experience abroad		
Erasmus experience	0.9245	0.8227*
Other Study experiences abroad	1.205	1.0545
Missing	0.9828	0.8749*
Post-graduate studies		
Post-graduate studies or professional experiences		
Training, apprenticeship aimed at gaining access to a liberal profession	0.5342***	0.5280***
Doctoral studies	0.7701	0.8794
Specialisation school	0.4980***	0.4930***
Ist level master degree	0.7819**	0.9925
IInd level Master degree	0.6551***	0.7579***
Other type of Master degree	0.9797	1.0055
Stage / Work grant / Training on-the-job	0.8862	1.0935
Public off-the-job training scheme	0.9378	0.9424
Study scholarship	0.8107	0.7053**
Voluntary civil service	1.2048	1.149
Movers and stayers		
Default: S(he) has not moved from the South and Islands		
Not moved from the North-Western Regions	0.7283***	0.7428***
Not moved from the North-Eastern Regions	0.7059***	0.7436***
Not moved from the Central Regions	0.9053	1.0346
Moved within the Northern Regions	0.6776**	0.8608
Moved within the Central Regions	(omitted)	(omitted)
Moved within the Southern and Islands Regions	1.1731	2.0061***
Moved to the North-West Regions	0.6593***	0.9224
Moved to the North-Eastern Regions	0.7215**	0.9144
Moved to the Central Regions	0.8151	0.9153
Moved to the Southern and Islands Regions	1.7623	1.1552
Moved abroad	0.6740**	0.7722
Constant	0.0103***	0.0182***
Number of observations	17387	17387
Pseudo R^2	0.1066	0.0989
Correctly classified cases	92.01%	88.57%
Area under the ROC curve	0.7493	0.7315

*Note: Legend: * p<0.05; ** p<0.01; *** p<0.001.*

The figures in the Table represent odds ratios. The odds ratio associated to a characteristic j is the relative risk of being overeducated for individuals with a given characteristics with respect to the reference or default group. E.g., if the estimated odds ratio of a given characteristics, say being IV years late in getting the degree equals 1.5, the corresponding group of graduates have a 50% higher probability of experiencing overeducation 5 years from the degree than the reference group of graduates who graduated in time; if the odds ratio equals 0.5 the individual with characteristics j have 50% lower probability of experiencing overeducation than the reference group.

Source: own elaboration on AlmaLaurea data.

Table 3. The wage penalty of overeducation and overskilling

	Overeducation (to get)	*Overskilling (to do)*
Dependent variable: Natural logarithm of net monthly wages	***(1)***	***(2)***
Unconditional estimates		
OLS	-0,2081***	-0,1568***
Interval regression	-0,2463***	-0,2088***
Conditional estimates		
OLS	-0,1220***	-0,0692***
Interval regression	-0.1319***	-0.0967***
Number of observations	16591	16591
Controlling for sample selection bias	Without instrumental variables	
Heckman model (ML simultaneous)	-0,1335***	-0,0758***
Heckman model (two steps)	-0,1336***	-0,0758***
	With instrumental variables (parents' education)	
Heckman model (ML simultaneous)	-0.1225***	-0,0758***
Heckman model (two steps)	-0,1337***	-0,0759***
Number of observations	21605	21605

Note: The table reports only the coefficients of interest. The OLS conditional estimates are obtained with all the control variables included in Table 4. The Heckit based on Maximum Likelihood simultaneous estimate are obtained with all the control variables included in Table 6. The two step estimates are unreported.
*Legend: * p<0.05; ** p<0.01; *** p<0.001.*
Source: own elaboration on AlmaLaurea data.

Table 4. OLS earnings equation augmented for overeducation or overskilling

Dependent Variable:	Log of net monthly wage	
Overeducation (to get)	-0,1220***	
Overskilling (to do)		-0.0692***
Gender. Default: Men		
Donna	-0.1849***	-0.1858***
Civil Status. Default: Single		
Married	0.1321***	0.1260***
Lives together	0.1071**	0.1064**
Separated, divorced, widow	0.0885	0.0565
Civil Status. Default: Single		
Married	-0.0064	-0.0043
Lives together	0.0718***	0.0688**
Separated, divorced, widow	0.2378***	0.2358***
Number of children, Men	0.1567***	0.1432***
Number of children, Women	0.0964***	0.0935***
Non Italian	-0.0508**	-0.0590**
Type of Secondary high school diploma. Default: Lyceum in sciences		
Classical high school	-0.0016	-0.0062
Specialisation in teacher training	0.0619***	0.0608***
Language high school	0.0081	0.0026
Art school	-0.0365	-0.0443
Technical school	-0.0103	-0.0072
Professional school	0.0178	-0.0101
Other high school diploma	-0.0042	-0.0104
Final grade at the university. Default: Magna cum laude		
66-90 out of 110	-0.0865***	-0.0770***
91-100 out of 110	-0.0598***	-0.0524***
101-105 out of 110	-0.0257**	-0.0187*

106-110 out of 110	-0.0235**	-0.0171*
Field of study. Default: Engineering		
Agriculture	-0.2958***	-0.2476***
Economics and Statistics	-0.2750***	-0.2327***
Law	-0.1231***	-0.0853***
Education	-0.4668***	-0.4136***
Atrts	-0.2430***	-0.1972***
Languages	-0.2817***	-0.2215***
Political and Social Sciences	-0.3353***	-0.2866***
Mathematics and Physics	-0.4171***	-0.3732***
Time to get a degree. Default: Curricular years		
I extra-curricular year late	-0.0726***	-0.0260*
II extra-curricular year late	-0.0969***	-0.0490***
III extra-curricular year late	-0.0957***	-0.0472***
IV extra-curricular year late	-0.1232***	-0.0754***
V extra-curricular year late	-0.1281***	-0.0825***
Study abroad. Default: No study experience abroad		
Erasmus experience	0.0188	0.0186
Other Study experiences abroad	0.0523***	0.0471**
No work experience	0.0721***	0.0695***
Post-graduate studies or professional experiences		
Training, apprenticeship aimed at gaining access to a liberal profession	-0.0876***	-0.1151***
Doctoral studies	-0.1050***	-0.0914***
Specialisation school	0.0065	-0.0262**
Ist level master degree	-0.0083	-0.0029
IInd level Master degree	-0.0062	-0.0032
Other type of Master degree	0.0068	0.0051
Stage / Work grant / Training on-the-job	0.0177**	0.0238***
Public off-the-job training scheme	-0.0550***	-0.0563***
Study scholarship	-0.0194	-0.0163
Voluntary civil service	-0.1081***	-0.1095***
Stayer versus movers. Default: S(he) has not moved from the South and Islands		
Not moved from the North-Western Regions	0.1580***	0.1562***
Not moved from the North-Eastern Regions	0.1535***	0.1504***
Not moved from the Central Regions	0.1257***	0.1216***
Moved within the Northern Regions	0.2433***	0.2421***
Moved within the Central Regions	(omitted)	(omitted)
Moved within the Southern and Islands Regions	0.1184***	0.1157***
Moved to the North-West Regions	0.2619***	0.2675***
Moved to the North-Eastern Regions	0.1793***	0.1795***
Moved to the Central Regions	0.1331***	0.1335***
Moved to the Southern and Islands Regions	0.0826	0.0845
Moved abroad	0.4730***	0.4703***
Constant	7.4377***	7.3524***
Number of observations	16591	16591

Note: The estimates regard a cross-section of pure pre-reform students who graduated in 2005, observed in 2010.
*OLS estimates. Legend: * p<0.05; ** p<0.01; *** p<0.001.*
Source: own elaboration on AlmaLaurea data.

Table 5. Different specifications of the wage penalty

	OLS unconditional estimates	OLS conditional estimates	IR unconditional estimates	IR conditional estimates	Heckit (no instruments)	Heckit (with instruments)
Overeducation only (to get)	-0.1629***	-0.1101***	-0.1832***	-0.1126***	-0.1182***	-0.1207***
Overskilling only (to do)	-0.0935***	-0.0179	-0.1470***	-0.0517***	-0.0193	-0.0230***
Overeducation and overskilling	-0.2392***	-0.1305***	-0.2910***	-0.1603***	-0.1440***	-0.1491***

Note: The table reports the coefficient of overeducation variables, with a distinction between those who are only overeducated, only overskilled or both overeducated and overskilled. The dependent variable is the natural logarithm of the net monthly wage. Conditional estimates are obtained with all the control variables included in Table 4. The Heckit estimates with no instruments are the same as in Table 6. those with instruments use the educational level of fathers and mothers as instrumental variables.
*Legend: * p<0.05; ** p<0.01; *** p<0.001.*
Source: own elaboration on AlmaLaurea data.

Table 6. Earnings equations with correction for sample selection

	Main equation	
Dependent Variable:	**Log of net monthly wage**	
Overeducation (to get)	-0.1225***	
Overskilling (to do)		-0.0758***
Gender. Default: Men		
Donna	-0.1547***	-0.1834***
Civil Status. Default: Single		
Non-single	0.1344***	0.1211***
Civil Status. Default: Single		
Non-single	0.0557***	0.0225*
Number of children, Men	0.1336***	0.1321***
Number of children, Women	0.0870***	0.0665***
Non Italian	0.0321	0.0041
Type of Secondary high school diploma. Default: Lyceum in sciences		
Classical high school	0.0152	-0.0078
Specialisation in teacher training	0.0504***	0.0497***
Language high school	0.0119	0.0036
Art school	0.0007	-0.0481
Technical school	-0.0138	-0.0133
Professional school	0.0505*	-0.0124
Other high school diploma	0.0204	-0.0194
Final grade at the university. Default: Magna cum laude		
66-90 out of 110	-0.0431**	-0.0635***
91-100 out of 110	-0.0564***	-0.0469***
101-105 out of 110	-0.0373***	-0.0207*
106-110 out of 110	-0.0493***	-0.0197*
Field of study. Default: Engineering		
Agriculture	-0.3152***	-0.2617***
Architecture	-0.2924***	-0.2403***
Economics and Statistics	-0.1854***	-0.0956***
Physical Education	-0.4398***	-0.4149***
Geology and biology	-0.1751***	-0.1987***
Law	-0.2893***	-0.2369***

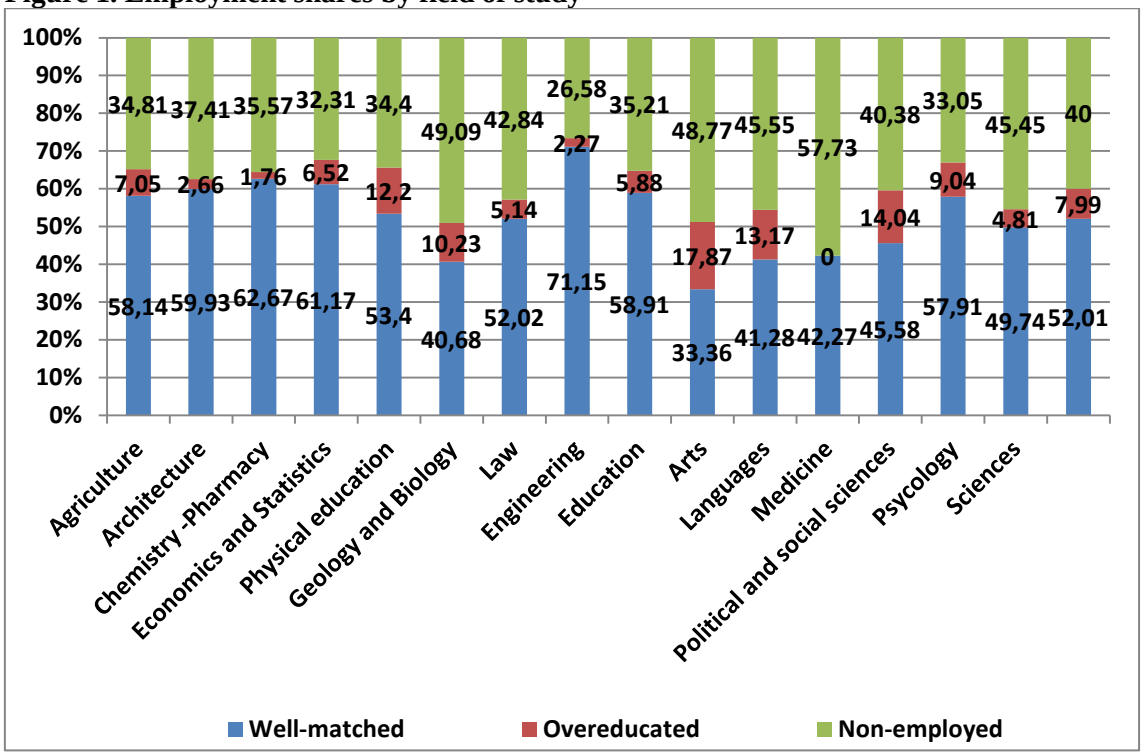
Education	-0.3832***	-0.2995***
Arts	-0.3365***	-0.3680***
Languages	-0.2613***	-0.2692***
Political and Social Sciences	-0.2562***	-0.1929***
Mathematics and Physics	-0.4767***	-0.3683***
Time to get a degree. Default: Curricular years		
I extra-curricular year late	-0.1032***	-0.0133
II extra-curricular year late	-0.1434***	-0.0441***
III extra-curricular year late	-0.1536***	-0.0481***
IV extra-curricular year late	-0.1692***	-0.0792***
V extra-curricular year late	-0.1670***	-0.0941***
Study abroad. Default: No study experience abroad		
Erasmus experience	0.0571***	0.0696***
Other Study experiences abroad	0.0617***	0.0648***
No work experience	0.0377***	0.0536***
Post-graduate studies or professional experiences		
Training, apprenticeship aimed at gaining access to a liberal profession	-0.0674***	-0.1375***
Doctoral studies	0.1143***	-0.0622**
Specialisation school	0.0186	-0.0188
Ist level master degree	-0.0213	-0.0016
IInd level Master degree	0.0065	0.0105
Other type of Master degree	0.0207	0.0168
Stage / Work grant / Training on-the-job	-0.0112	0.0252***
Public off-the-job training scheme	-0.0494***	-0.0640***
Study scholarship	0.0239	0.0047
Voluntary civil service	-0.1008***	-0.1285***
Stayer versus movers. Default: S(he) has not moved from the South and Islands		
Not moved from the North-Western Regions	-0.0994***	0.0696***
Not moved from the North-Eastern Regions	-0.1146***	0.0626***
Constant	7.7484***	7.4408***
Selection equation		
Dependent Variable:	Being employed	
Gender. Default: Men		
Donna	-0.1401***	-0.1369***
Civil Status. Default: Single		
Non-single	-0.0496	-0.0671
Civil Status. Default: Single		
Non-single	-0.1201***	-0.1592***
Number of children, Men	0.1031	0.1207
Number of children, Women	-0.0968*	-0.0895
Non Italian	-0.0058	-0.0259
Type of Secondary high school diploma. Default: Lyceum in sciences		
Classical high school	-0.0656**	-0.0575**
Specialisation in teacher training	-0.047	-0.0002
Language high school	0.0061	0.0082
Art school	-0.1476*	-0.1979**
Technical school	0.006	-0.0299
Professional school	-0.1499**	-0.1083
Other high school diploma	-0.1652**	-0.1785**
Final grade at the university. Default: Magna cum laude		
66-90 out of 110	-0.1151**	-0.2063***
91-100 out of 110	-0.0036	-0.0472
101-105 out of 110	0.0233	-0.0065
106-110 out of 110	0.0788***	0.0705**
Field of study. Default: Engineering		
Agriculture	0.1248*	-0.2952***
Architecture	0.1490***	-0.2296***
Economics and Statistics	0.2108***	-0.0308

Physical Education	-0.0112	-0.5050***
Geology and biology	-0.2395***	-0.6327***
Law	0.1018***	-0.4156***
Education	0.1144**	-0.1760***
Arts	-0.2256***	-0.6716***
Languages	-0.1730***	-0.5688***
Political and Social Sciences	0.1323***	-0.1980***
Mathematics and Physics	0.2920***	-0.2277***
Time to get a degree. Default: Curricular years		
I extra-curricular year late	0.1930***	-0.0315
II extra-curricular year late	0.2321***	-0.0038
III extra-curricular year late	0.2501***	0.0213
IV extra-curricular year late	0.1874***	-0.0639
V extra-curricular year late	0.1679***	-0.1155**
Study abroad. Default: No study experience abroad		
Erasmus experience	0.1189***	0.0949**
Other Study experiences abroad	0.1059**	0.1012*
No work experience	0.0586	0.0879
Post-graduate studies or professional experiences		
Training, apprenticeship aimed at gaining access to a liberal profession	-0.0398	-0.0151
Doctoral studies	-0.5771***	-0.8161***
Specialisation school	0.0531*	0.0766**
Ist level master degree	0.0389	0.0408
IInd level Master degree	0.0024	-0.0212
Other type of Master degree	0.041	-0.0055
Stage / Work grant / Training on-the-job	0.1053***	0.1194***
Public off-the-job training scheme	-0.0466	-0.0670**
Study scholarship	-0.0899*	-0.1359**
Voluntary civil service	-0.1148**	-0.1471***
Stayer versus movers. Default: S(he) has not moved from the South and Islands		
Not moved from the North-Western Regions	0.9011***	1.1538***
Not moved from the North-Eastern Regions	0.9193***	1.2082***
Father's education. Default: Compulsory or below		
High secondary school	0.0395*	0.0354
University	0.0298	0.0441
Don't know	-0.0395	-0.041
Mother's education. Default: Compulsory or below		
High secondary school	0.0781***	0.0614**
University	0.0577*	0.0209
Don't know	0.1400*	0.1087
Constant	0.3824***	1.0523***
Arthro	-1.3342***	-0.0203*
Lnsigma	-0.6361***	-0.8066***
Number of graduates who declare their wage	16591	16591
Total number of graduates	21605	21605

Note: Pure pre-reform graduates in 2005, observed 5 years after graduation. ML simultaneous estimate.
** p<0.05; ** p<0.01; *** p<0.001.*
Source: own elaboration on AlmaLaurea data.

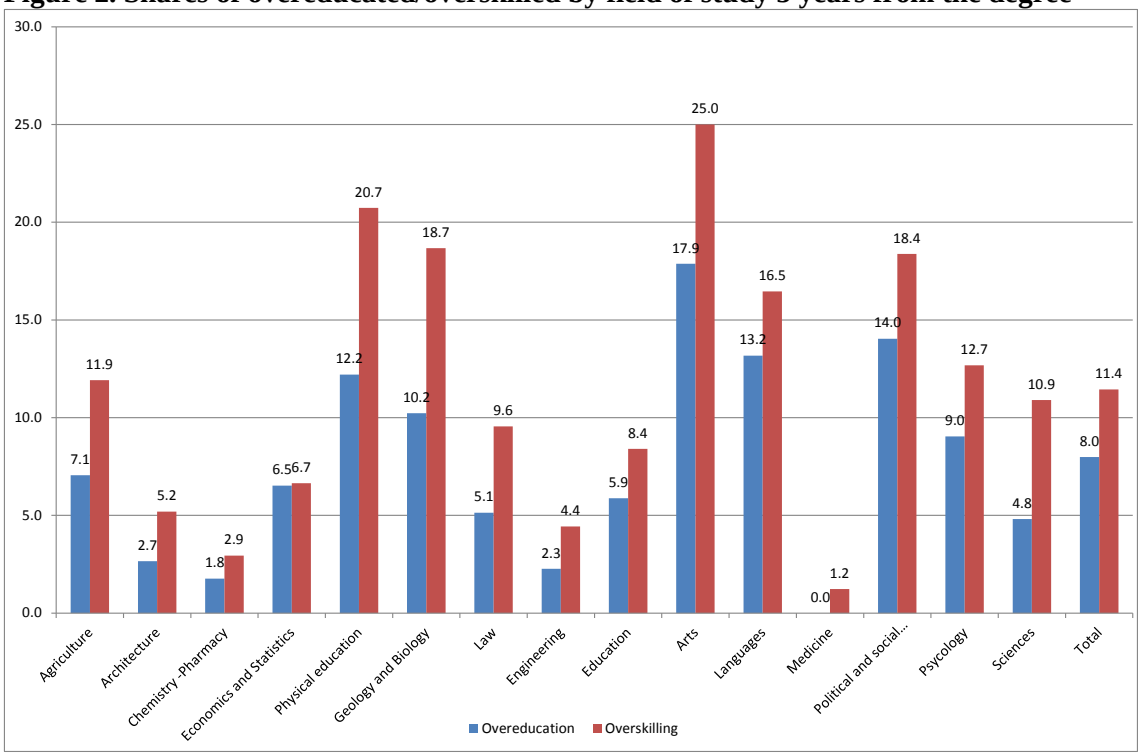
Figures

Figure 1. Employment shares by field of study



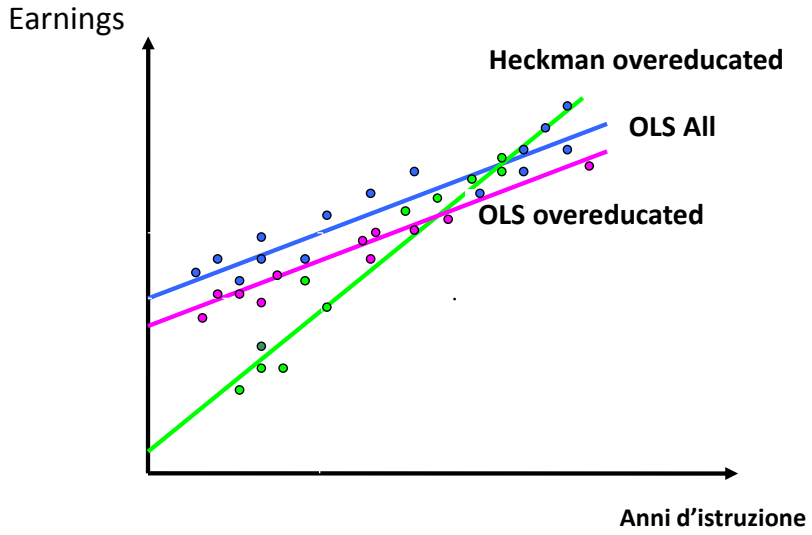
Source: own elaboration on AlmaLaurea data.

Figure 2. Shares of overeducated/overskilled by field of study 5 years from the degree

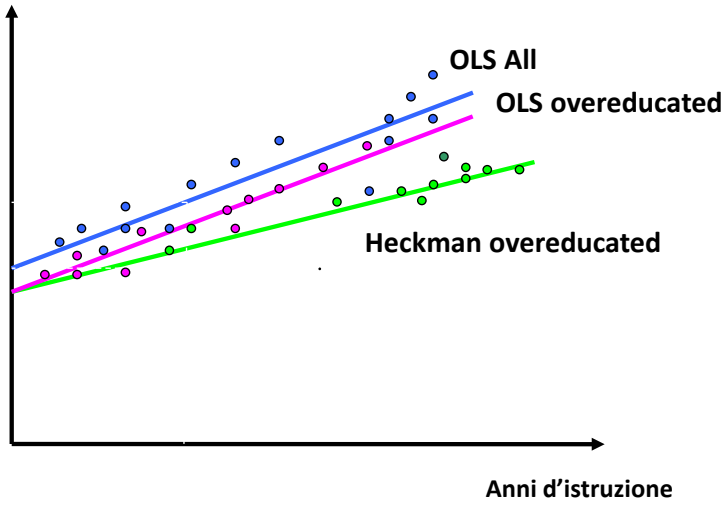


Source: own elaboration on AlmaLaurea data.

Figure 3. Heckman correction of the wage effect of overeducation/overskilling
Panel (a): Job competition, job assignment and human capital model



Panel (b): Job search model



Annex

Table A.1. Descriptive statistics (mean and shares) for pre-reform graduates 5 years after graduation

Variable	*Total employment*	*Overeducated*	*Non-Overeducated*	*Overskilled*	*Non-overskilled*
Overeducated	0.0799	1	0	0.4686	0.0297
Overskilled	0.1144	0.6710	0.0661	1	0
Donne	0.5974	0.6278	0.5948	0.6611	0.5892
Civil status. Default: Single (men)					
Married	0.0270	0.0324	0.0266	0.0261	0.0271
Living together	0.0053	0.0065	0.0053	0.0050	0.0054
Separated, divorced, widow	0.0012	0.0007	0.0013	0.0010	0.0012
Civil status. Default: Single (women)					
Married	0.0671	0.0778	0.0661	0.0814	0.0652
Living together	0.0130	0.0166	0.0127	0.0156	0.0127
Separated, divorced, widow	0.0037	0.0043	0.0036	0.0025	0.0038
Sons (men)	0.0180	0.0173	0.0181	0.0156	0.0183
Sons (women)	0.0345	0.0410	0.0339	0.0452	0.0331
Foreigners	0.0252	0.0274	0.0250	0.0261	0.0251
Type of high secondary school diploma. Default: Second Level College of Science					
Secondary school focusing on humanities	0.2067	0.1944	0.2078	0.2227	0.2046
Secondary school focusing on pedagogy	0.0724	0.0684	0.0728	0.0835	0.0710
Secondary school focusing on languages	0.0546	0.0886	0.0516	0.0749	0.0520
Artistic high school	0.0152	0.0295	0.0140	0.0251	0.0140
Technical high school	0.2244	0.2527	0.2220	0.2167	0.2254
Professional high school	0.0234	0.0360	0.0223	0.0362	0.0217
Other	0.0127	0.0187	0.0122	0.0161	0.0123
Final grade at the degree. Default: summa cum laude					
66-90	0.0703	0.0713	0.0703	0.0729	0.0700
91-100	0.2751	0.3017	0.2728	0.2846	0.2739
101-105	0.2206	0.2253	0.2202	0.2433	0.2176
106-110	0.2314	0.2505	0.2297	0.2524	0.2287
Field of study. Default: Engineering					
Agricolture	0.0212	0.0187	0.0214	0.0221	0.0211
Architecture	0.0476	0.0158	0.0504	0.0216	0.0510
Economics and Statistics	0.1367	0.1116	0.1388	0.0794	0.1440
Physical education	0.0047	0.0072	0.0045	0.0085	0.0042
Geology and Biology	0.0320	0.0410	0.0313	0.0523	0.0294
Law	0.1265	0.0814	0.1305	0.1056	0.1292
Pedagogy	0.0753	0.0554	0.0771	0.0553	0.0779
Arts	0.1081	0.2419	0.0965	0.2363	0.0916
Languages	0.0594	0.0979	0.0561	0.0855	0.0560
Medicine	0.0374	0.0000	0.0407	0.0040	0.0418
Political Science	0.1217	0.2138	0.1137	0.1956	0.1122
Psycology	0.0553	0.0626	0.0547	0.0613	0.0546
Mathematics and phisics	0.0179	0.0108	0.0186	0.0171	0.0181
Years "fuoricorso". Default: degree in corso					
I year delay	0.2188	0.1728	0.2228	0.1800	0.2238
II year delay	0.1989	0.1908	0.1997	0.1785	0.2016
III year delay	0.1525	0.1512	0.1526	0.1644	0.1509

IV year delay	0.0994	0.1044	0.0990	0.1091	0.0982
V year delay	0.2609	0.3492	0.2532	0.3333	0.2515
Study abroad. Default: No study experience abroad					
Erasmus	0.0661	0.0641	0.0663	0.0593	0.0670
Other types of experience abroad	0.0381	0.0533	0.0368	0.0463	0.0371
Omitted	0.1571	0.1800	0.1551	0.1659	0.1560
Post-graduate studies or professional experiences					
Training, apprenticeship aimed at gaining access to a liberal profession	0.2791	0.1641	0.2890	0.1760	0.2924
Doctoral studies	0.0274	0.0158	0.0284	0.0186	0.0286
Specialisation school	0.1142	0.0583	0.1190	0.0649	0.1205
Ist level master degree	0.0784	0.0713	0.0791	0.0870	0.0773
IInd level Master degree	0.0805	0.0605	0.0822	0.0684	0.0820
Other type of Master degree	0.1008	0.1058	0.1004	0.1066	0.1001
Stage / Work grant / Training on-the-job	0.2065	0.2009	0.2070	0.2272	0.2038
Public off-the-job training scheme	0.1215	0.1353	0.1203	0.1368	0.1195
Study scholarship	0.0326	0.0194	0.0337	0.0196	0.0342
Voluntary civil service	0.0378	0.0533	0.0364	0.0523	0.0359
Movers and stayers. Default: He has not moved from Mezzogiorno and Islands					
***He has not moved from* North-West**	0.1280	0.1159	0.1291	0.112	0.1301
***He has not moved from* North-East**	0.1942	0.1785	0.1955	0.178	0.1962
***He has not moved from* Centre**	0.0095	0.0101	0.0095	0.010	0.0095
He has moved within the Northern area	0.0369	0.0302	0.0375	0.035	0.0372
He has moved within the Central area	0.0095	0.0101	0.0095	0.010	0.0095
He has moved within the Southern area and islands	0.0083	0.0094	0.0082	0.013	0.0077
He has moved towards the North-West	0.0524	0.0367	0.0538	0.048	0.0530
He has moved towards the North-East	0.0414	0.0346	0.0419	0.041	0.0414
He has moved towards the Centre	0.0565	0.0540	0.0567	0.057	0.0564
He has moved towards the South	0.0054	0.0086	0.0051	0.006	0.0053
He has moved abroad	0.0311	0.0274	0.0314	0.028	0.0315
Total employment	**17387**	**1389**	**15998**	**1989**	**15398**

Notes: 1 I level master means a master program that can be accessed after a 3-year program; a II level master means a master program that can be accessed after a 3+2 program.
Source: own elaboration on AlmaLaurea data.