Since 1994 Inter-University Consortium
AlmaLaurea
Connecting Universities, the Labour Market and Professionals

AlmaLaurea Working Papers – ISSN 2239-9453

ALMALAUREA WORKING PAPERS no. 77

December 2016

Overeducation among Italian graduates: do different measures actually diverge?

by

Luca Cattani[1], Giovanni Guidetti[2], Giulio Pedrini[3]

[1]*Department of Economics and School of Development Innovation and Change (SDIC), University of Bologna.*

[2]*Department of Economics and School of Development Innovation and Change (SDIC), University of Bologna.*

[3]*Interuniversity Research Centre for Public Services (CRISP), University of Milan-Bicocca,*

School of Development Innovation and Change (SDIC), University of Bologna

The **AlmaLaurea working paper series** is designed to make available to a wide readership selected works by AlmaLaurea staff or by outside, generally available in English or Italian. The series focuses on the study of the relationship between educational systems, society and economy, the quality of educational process, the demand and supply of education, the human capital accumulation, the structure and working of the labour markets, the assessment of educational policies.

Comments on this series are welcome and should be sent to pubblicazioni@almalaurea.it.

AlmaLaurea is a public consortium of Italian universities which, with the support of the Ministry of Education, meets the information needs of graduates, universities and the business community. AlmaLaurea has been set up in 1994 following an initiative of the Statistical Observatory of the University of Bologna. It supplies reliable and timely data on the effectiveness and efficiency of the higher education system to member universities' governing bodies, assessment units and committees responsible for teaching activities and career guidance.

AlmaLaurea:

- facilitates and improves the hiring of young graduates in the labour markets both at the national and international level;
- simplifies companies' search for personnel, reducing the gap between the demand for and supply of qualified labour (www.almalaurea.it/en/aziende/);
- makes available online more than 1.5 million curricula (in Italian and English) of graduates, including those with a pluriannual work experience (www.almalaurea.it/en/);
- ensures the optimization of human resources utilization through a steady updating of data on the careers of students holding a degree (www.almalaurea.it/en/lau/).

Each year AlmaLaurea plans two main conferences (www.almalaurea.it/en/informa/news) in which the results of the annual surveys on Graduates' Employment Conditions and Graduates' Profile are presented.

AlmaLaurea Inter-University Consortium | viale Masini 36 | 40126 Bologna (Italy)

Website: www.almalaurea.it | E-mail: pubblicazioni@almalaurea.it

Overeducation among Italian graduates: do different measures actually diverge?

LUCA CATTANI[1], GIOVANNI GUIDETTI[2], GIULIO PEDRINI[3]

Abstract

In this paper, we explore three dimensions of educational mismatch among graduates: incidence, impact on earnings and possible determinants of overeducation. Our analysis focuses on Italian graduates and refers to the cohort that graduated in 2007 using data from the AlmaLaurea survey on graduates entering the labour market. A new measure of overeducation is introduced and jointly examined along with an alternative measure based on workers' self-assessment. After having run estimates of the impact of overeducation on earnings and analyzed possible determinants of educational mismatch, we conclude that the two definitions of overeducation measure quite different things and in particular that "traditional" measures based on workers' self-assessment are affected by individuals' characteristics and by workers' expectations and perceptions concerning the job post. However, effects on wages are very similar no matter what definition is adopted.

JEL classification: I2, J31

Keywords: educational mismatches, human capital, graduate labour markets.

Acknowledgments

This paper has been written relying on elaborations on AlmaLaurea data: we are grateful to Andrea Cammelli, Angelo Di Francia and Silvia Ghiselli for their support and their useful comments and suggestions. We have also utilised data from the survey on Italian professions: elaborations on these data have been run under our responsibility. Usual disclaimers apply.

[1] *Department of Economics, SDIC, University of Bologna*

[2] *Department of Economics, SDIC, University of Bologna*

[3] *Interuniversity Research Centre for Public Services (CRISP), University of Milan-Bicocca, SDIC, University of Bologna*

1 Introduction

Overeducated workers are basically individuals endowed with educational attainments, were their knowledge, competences or skills, in excess of what is actually needed or required to perform tasks associated to their current job. The economic literature on overeducation starts with Freeman (1976) as an aggregate study on decreasing returns to education investments, proxied by the average college premium paid to graduate workers in American labor markets. In this view, public and private overinvestments in education result in lower levels of returns due to the fact that the supply of highly qualified labour is outpacing its relative demand and thus causing a depreciation of college premiums. At a micro level, overeducation is interpreted as a source of inequality among peers, such as workers with the same educational levels but earning different wages once employed in differently demanding jobs (Frank, 1978; Berg, 1970). Duncan and Hoffman (1981) implement an extended version of the Mincerian equation in order to estimate separately the effects on wages of required, surplus or deficit years of schooling and kick-start the overeducation literature, a popular and much debated economic subfield lying in between labour economics and the economics of education. At the operational level, measuring overeducation consists in assessing the gap between the required and the attained years of schooling for each individual in a given sample. However, while it is quite easy to assess employees' education with a specific additional question in the questionnaire, measuring what employers are effectively demanding has proved to be less simple, dividing the most of the contributions to this debate between supporters of workers' self-assessment (WA) or job analysis (JA) measures. On the one side, WA measures can be severely biased by individuals' characteristics and perceptions. On the other side, JA may neglect the potential influence of supply forces. We contribute to this debate by introducing a new measure for job requirements based on the classification of occupations proposed by Cattani et al. (2014) and assessing its validity with respect to Italian graduates. With respect to the existing literature, the new measure allows to fully exploit the information concerning the cognitive and skill content of the tasks associated to a particular job in a precise and up-to-date manner. Our purpose is to determine the incidence, determinants and impact on earnings of overeducation by using this new measure and iterate the same analysis utilizing an alternative measure based on workers' self-assessment in order to compare the two different outcomes and to show the benefits of the new measure. The Italian context represents an interesting case study. In early 2000s Italy has experienced a sudden increase in the number of graduates due to the participation expansion in tertiary education. In addition, Italy has completed in this decade the implementation of the so-called "3+2" system[4]. This expansion combined with the dramatic recession that hit the country in the 2008-2012 period, has raised growing concerns on graduates' overall employability and wage penalties associated with overeducation.

The structure of the paper is as follows. Paragraph §1 introduces the most relevant measurement issues, paragraph §2 presents the new measurement of overeducation adopted in the paper, paragraphs §3 and §4 describe the estimation methodology along with the dataset, paragraphs §5 presents and discusses estimates concerning determinants of overeducation and its impact on wages. Conclusions will follow.

[4] The reform is termed "3+2" and implements the so-called "Bologna process" being it based onto a two-cycle degree structure: a first-level three-year undergraduate course plus a second-level two-year master degree. A few programmes maintained their five/six-year single-cycle structure.

2 The measurement of overeducation in the economic literature

Educational attainments represent only the supply side in the matching between demand and supply of qualified labour. Measuring workers' titles and skills therefore constitutes only half of the work we should do in order to assess the overeducation incidence and its relative wage effects. The education and skill levels needed to competently perform the constituent tasks of a particular job are, in fact, the benchmark to which we have to refer individuals' bundle of human capital in order to understand who is matched, who has deficit and who has excess schooling. Up to now, the economic literature has adopted three main methods for proxying for job requirements.

The first and most utilized strategy to measure job requirements is to directly ask workers what is actually required or needed to obtain or carry out the job, i.e. a worker self-assessment (WA measure). The seminal paper of Duncan and Hoffman (1981), along with others (Hartog and Tsang, 1987; Sicherman, 1991), refer to the WA formal education required to *obtain* the job, while Ramirez (1993) referred to the WA informal education needed to *perform* the job. Nonetheless, as WA reflects the worker's point of view it is suitable to bring to biases as workers could tend to overstate their job requirements to inflate their job position during the interview or, in newly hired workers, reflect qualification inflation in firms' hiring strategies (Hartog, 2000).

The other most common measurement is obtained by looking at information provided in the occupational classifications and thus building a correspondence table that assigns an educational title to each job title (JA measure). Many works adopted this strategy (Thurow and Lucas, 1972; Hartog, 1980; Rumberger, 1987; Kiker and Santos, 1991; Oosterbeek and Webbink, 1996) referring either to the General Educational Development (GED) taxonomy or to the Dictionary of Occupation Titles (DOT).

Unfortunately, this measurement didn’t gain an acceptable level of consensus as classifications are rarely updated because updates are costly/expensive (Mason, 1996; Hartog, 2000). Groot and Maassen van den Brink (2000) argue that each JA measure is only temporary valid and it should not be used to measure trends in overeducation. Moreover, operators in the field such as administrative staff in employment offices oppose frequent changes and try to defend their interests in the revision processes (Elias and McKnight, 2001) and there is no consensus when converting occupational classifications into schooling years (Halaby, 1994). However, according to Hartog (2000) JA is still conceptually superior to all other types of measurement as it is the only one with no expected biases.

Other authors have also looked at market realizations (RM measure) such as the mean educational attainment in a given occupation or hiring standards used by firms’ personnel departments (Verdugo and Verdugo, 1989; Groot and Maassen van der Brink, 1997; Groeneveld and Hartog, 2004). Accordingly, a job can be considered as a ‘graduate job’ on the basis of the mean college premium associated with it (Gottschalk and Hansen, 2003). Unfortunately, these matches are the result of demand and supply forces and don’t reflect job requirements only (Leuven and Oosterbeek, 2011). As a consequence, RM do not capture structural overeducation. When a given ‘non-graduate occupation’ is crowded with high proportions of overeducated workers, it is likely that it will be ranked as a ‘graduated job’ using RM thus leading to an underestimation of the overall overeducation incidence. Overall, in the economic literature RM measure is usually considered less reliable than JA and WA measures (Verhaest and Omey, 2006a)

Empirical studies using different measures of overeducation has reached ambiguous results. Although McGoldrick and Robst (1996) and Verhaest and Omey, (2006b) find that JA measures bring higher levels of overeducation than those assessed with the WA measure, in the vast majority of studies the incidence of JA measures is significantly lower than WA ones (McGuinness, 2006; Cedefop, 2010). Additional studies highlight how these two different approaches tend to generate poorly correlated measures (varying between 10% and 50%), with many individuals considered as overeducated under one definition and properly matched under the alternative one (Robst, 1994; Battu et al., 2000; Groot and Maassen van den Brink 2000). In terms of the determinants of overeducation, the use of JA and WA has produced quite different results, leading to the argument that "the various measures do not capture equally well overeducation or are likely to be indicators for related concepts" (Verhaest and Omey, 2010, p. 620). This divergence has been attributed to two sources of bias. On the one hand, the JA measure is affected by a certain degree of measurement error. On the other hand the WA measure is likely to take incorporate individuals' expectations and disappointments when assessing the educational mismatch (Verhaest and Omey, 2006a). Different outcomes could be also related to the heterogeneous degree of validity of JA measures and to the need of regularly updating them in order to correctly estimate the incidence and effects of overeducation (Van der Meer, 2006). Despite these differences, JA and WA measures lead to very similar estimates when considering the impact of overeducation on wages, suggesting the idea that estimates could actually capture unobserved individual variability, such as ability, correlated to both wages and overeducation (McGuinness, 2006). These biases have hindered a correct assessment of the actual differences between the different methods, leading to the frequent misperception that measurement issues are not significant in assessing the wage effects of overeducation. Quite the opposite, their implication is that measurement issues still need to be investigated at both conceptual and empirical levels (Leuven and Oosterbeek, 2011). This has called for the need to carry on further comparative analyses based on both JA and WA measures[5].

3 The new measure of overeducation adopted in the paper

In this study we infer job requirements from the Italian occupational classification (CP 2006) after having attached to each job title a European Qualification Framework (EQF)[6] corresponding level. In order to achieve this, we propose an application to the Italian labour force of the British SOC[7] (HE) 2010 classification first proposed by Elias and Purcell (2004; 2011) and then iterated and differently constructed by Cattani et al. (2014). In this iteration, termed SOC(HE)-Italy, the allocation of all 800 Italian occupational units to one of the constituent SOC(HE) Major Groups (Experts, Orchestrators, Communicators and Non-graduate-jobs) is based on data from the survey on Italian professions (ISTAT, 2009) in which 16,000 workers were asked to assign a score on a 1-100 point scale to 109 variables referred to the O*Net[8] taxonomy for knowledge, skills and competences utilized on the job[9].

[5] Interestingly, Robst (1994) shows in a comparative study how, when instrumented, estimates of the wage penalty associated with JA measures tend to be more stable than those obtained with WA measures. This fact points out how these last ones could be more biased

[6] The European Qualification Framework (EQF) is a common transnational translation device for all European qualifications. Qualifications are here defined as educational titles issued at the completion of an educational or training process. The aim of the EQF (issued by the European Commission in 2008) is to make different national qualifications more readable across the continent and "promoting workers' and learners' mobility between countries and facilitating their lifelong learning" (Recommendation 111/2008). It relates all European national qualifications to 8 major levels, referring to knowledges, skills and competences acquired in their relative education/training process. In our study, this is of crucial importance given the univocal translation from Italian qualifications into EQF levels letting room for a univocal translation of EQF levels into schooling years.

[7] SOC stands for British Standard Occupational Classification.

[8] O*Net (Occupational Information Network) is an American data collection and spreading system focused on employment, jobs, skills shortages, professional profiles and individual characteristics. It is based on the SOC classification and it has been structured to describe tasks and professional profiles demanded and supplied enacting work processes. O*Net embodies the advantages of SOC classification and its implementation took large account of the indications emerged from the SCAN works, such as the

The allocating procedure is as follows:

First, a mean difficulty score of the 109 variables on the 1-100 point scale provided by the questionnaire is calculated for each of the 800 Italian job titles.

Second, only those variables exceeding a standard deviation from the mean difficulty score are selected for each job title.

Third, the selected variables are regrouped into the three clusters of competence identified by Purcell et al. (2012), which correspond in turn to the three SOC(HE) Major Groups of Graduate Professions:

a. *Specialist expertise deriving from HE knowledge*. Consisting in "detailed knowledge and/or skills for which the normal foundation is an undergraduate degree course and where these are continually being exercised, developed and/or refined in practical and/or theoretical terms";

b. *Orchestration expertise*. Consisting in "high-level competence based on knowledge and skills that may have been developed either in HE or through experience (and most often, both of these). It incorporates the ability to draw together knowledge and knowledge-holders, to direct and co-ordinate activities, assess alternatives,

---

distinction within the three types (basic, thinking and personal) of *soft skills*. It is divided into six dimensions: *Experience Requirements*, *Occupation Requirements*, *Occupation Specific Information*, *Occupation Characteristics*, *Worker Characteristics* and *Worker Requirements*. This particular structure allows the in-depth description of different job profiles and it is fit, thanks to transcode tools, to networking by exploiting linkages with other classification systems.

[9] Respondents to the Istat survey were asked to assign a score to 255 variables on a 1-100 point scale in terms of complexity of the knowledge, skill or competence associated with the tasks of their particular job, making thus reference to the level of knowledge and skills rather than to the their frequency of use. These 255 variables were borrowed from the O*Net6 taxonomy and covered 7 areas: Knowledge (33 questions), Skills (35 questions), Attitudes (52 questions), Values (21 questions), Working styles (16 questions), Generalized working activities (41 questions) and Working conditions (57 questions). Starting from this point, we analyzed only the 109 variables contained in the three areas referred to knowledge, skills and tasks/competences ('Knowledge', 'Skills' and 'Generalized working activities') as described later on in this paragraph. Questions in the Istat survey were formulated in the same way for all the 109 variables, as follows: "At what level, among those listed [1-100], you need this knowledge/skill/activity to fulfill your current profession?". Variables in the 'knowledge' section included, for istance, 'B.3 - Economics and accounting: knowledge of the principles and methods of accounting, financial markets, and analyzing and dissemination techniques of financial data'. The full list of variables included in the survey is available at http://fabbisogni.isfol.it/

evaluate risks and influence or make high-level decisions on the basis of available evidence”;

c. *Communication expertise*. “Consisting in knowledge and skills, normally involving well-developed interactive skills, concerned with the exercise of high-level competence in the communication, dissemination and use of knowledge, ideas and information, between individuals, within groups, or for mass-production or consumption, delivered in person or using digital media.” (Elias and Purcell, 2011)

Fourth, three difficulty scores are computed for each of the 800 Italian job titles which simply equal the mean of the variables on the 1-100 point scale included in the three clusters.

Fifth, the 1-100 point scale difficulty scores are translated into the 1-8 point scale of the EQF, following two simple assumptions.

- a linear progression of returns to education in terms of knowledge and skills is postulated. This assumption is mutuated from ISFOL’s “Navigatore delle professioni”[10]. Consistently, a simple proportion between the two scales is run:

(1) $$EQF_i = \left(\frac{DS_i}{100}\right) \times \mathbf{8}$$

  Where EQF_i is the resulting EQF level of the job title “i” and DS_i is its relative difficulty score;

- a double cut-off is then applied in order to discriminate between ‘graduate’ and ‘non-graduate’ jobs: 5.5 for EQF_i and 68.75 for DS_i :

[10] See http://fabbisogni.isfol.it//documenti/it/nup_eqf_2011.pdf

(2) $$\left[\left(\frac{5.5}{8}\right) * 100\right] = \mathbf{68.75}$$

This cut-off marks the boundary between 'graduate' and 'non-graduate' jobs. It is of course arbitrary and descends from a benchmarking procedure run against job analysis data utilizing the "Navigatore delle professioni" by ISFOL and the "Sistema informativo delle professioni" by ISTAT[11] applying alternative (slightly different) cut-offs.

Sixth, in case the job title has at least one difficulty score out of three exceeding the cut-off value, the job title is allocated into the Major Group corresponding to the cluster of competence that has recorded the highest score on the 1-8 point scale. In case none of the three score are above the threshold, the job title is allocated to the 'non-graduate jobs' Major Group.

Following this procedure economists are, for instance, assigned to Experts major group while their EQF level is 7, equivalent to the Italian Master's Degree (18 schooling years).

This job requirement measure shares with the JA approach the benefits of avoiding subjective respondents' bias driven by WA. In fact, job requirements are represented by educational requirements stated by workers without including job satisfaction dimensions and their subjective job position assessment. Interviews in the ISTAT survey are carried out referring explicitly only to skills, knowledge and competences utilization on the job place without mentioning job positions. A particular job is thus analyzed only referring to the cognitive content encompassed in its constituent tasks and duties and their relation to higher education outcomes, namely knowledge, skills and competences as systemized in the O*Net taxonomy. Moreover, workers are

[11] The Informative System on Italian Profession is available at:
http://professioni.istat.it/sistemainformativoprofessioni/cp2011/index.php?codice_1=2

sampled in the ISTAT survey on the basis of the position they hold in their organization and there is no room for them to overstate it. Finally, workers interviewed in ISTAT survey are not the same we observe in our sample of Italian graduates.

SOC(HE)-Italy measure for overeducation comes to be a sort of mixed JA measure expressed in schooling years which are in turn determined by EQF framework. This reference platform benefits from a certain degree of consensus, at least amongst decision makers, even if the translation from the original difficulty scores in the SOC(HE)-Italy into the corresponding EQF levels holds only under the hypothesis of linearity of returns to knowledge . On the other hand, JA measures are, as noted above, affected also by imprecision as they are costly to revise and thus rarely updated. Our measure could be, in other words, objective and precise to some extent but limited in time as professions evolve changing their typical tasks and their relative cognitive contents. Basing our measure on data from the ISTAT survey on Italian professions partially addresses this problem as this survey is periodically held by ISTAT. Therefore data availability should not represent a major problem with respect to Italy as the cognitive and skill content of the tasks associated to a particular job is taken into account in a precise and up-to-date manner. Overall, despite the complexity of the construction procedure of the SOC(HE) is counterbalanced by the regular revision of the data source.

4 The validity assessment of the new measure

In order to asses to what extent two measures actually diverge it is necessary to discuss their relative validity in capturing the ideas contained in the corresponding concepts

(Adcock and Collier, 2001). First, we need to assess three different types of validity: a) content validity; b) criterion validity and c) construct validity.

a) *Content validity* is based on the match between the “a priori” concepts developed in a theoretical framework and the indicators directly deriving from the concepts themselves. In this respect, for testing the content validity of a given measure, the first step consists in providing a systematized concept (i.e.: an explicit definition of a phenomenon to be analysed). A second step is represented by the operationalization of the systematized concept, that is the setting up of one or more indicators to score or classify the analysed cases. Then, both the indicators and the systematized concepts, from which the indicators stem, can be fine-tuned on the basis of insights deriving from a feedback process implying a continuous match between measure of the empirical phenomena and conceptual settings;

b) *Criterion validity* is based on the comparison between the scores produced by a certain indicator and the scores produced by other variables, considered a direct and reliable measure of the phenomenon of interest;

c) *Construct validity* is a process of validation aimed at assessing whether an indicator is linked to other indicators according to a framework which is consistent with the theoretical expectations. Substantially, an indicator is considered to be valid if the relations established between this indicator and other variables are consistent with a widely accepted and well-grounded theoretical proposition.

Then Adcock and Collier add to these three types of validation the so-called process of convergent/discriminant validation, which can be considered as a specific version of the criterion validity. Taking into account two alternative indicators which are supposed to measure the same systematized concept, one can refer to convergent validation when there is strong association among the indicators' scores. Conversely, when the scores of two alternative indicators exhibit a weak degree of association, then one can refer to discriminant validation and state that the two indicators are based on different systematized concepts.

As outlined in Cattani et al. (2014) the SOC(HE)-Italy classification is a valid indicator of the distinction between ‘graduates’ and ‘non-graduates’ jobs, being it based onto the linking between the skills acquired during the studies and the cognitive content encompassed in the constituent tasks and duties of a particular job. Consequently, the skill levels (in terms of EQF levels) attached to each job title by this classification enable us to construct a reliable measure of overeducation, based on a well-structured definition of the mismatch between labour demand and supply (e.g. job requirements and educational attainments) which is at the basis of this phenomenon.

Once acknowledged the content validation of SOC(HE)-Italy we need to assess the criterion validity and the convergence/discrimination validity between the SOC(HE)-Italy and an alternative WA measure through the estimate of their determinants. Then, the construct validation requires the existence of a well-established theoretical hypothesis, assumption or at least empirical finding which is widely accepted among the scholars in a certain topic. In this case we identify as generally accepted the idea that overeducated workers suffer from a wage penalty in comparison to their well-matched counterparts sharing the same educational attainments. The intuition behind this evidence is that an overeducated worker holds a job position in which s/he does not fully exploit his/her endowment of knowledge and skills as a worker with the same qualifications, but in a well matched job, would do.

5 The estimation methodology

As outlined in the previous subparagraphs, the new measure of overeducation is dichotomous and, therefore, its determinants can be estimated through a straightforward Probit model. Applying standard treatment of the Probit model, we have that Overeducation =1 (YES) when a latent variable Y is strictly positive, Y>0, and that Overeducation =0 (NO) when Y is nil, Y=0.

The latent variable is linked through a linear function to a set of statistical variables so that:

(1) $$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots . \beta_n X_n + \varepsilon_i = X'_i \beta + \varepsilon_i$$

where ε_i is a normally and independently distributed error term (NID). Consequently, we have:

(2) $$P(\textbf{Overeducation} = \textbf{1} = \textbf{YES}) = \textbf{P}(\textbf{Y} > \textbf{0}) = P(X'_i \beta + \varepsilon_i > \textbf{0}) = P(-\varepsilon_i \leq X'_i \beta) = F(X'_i \beta)$$

where F is the distribution function for ε_i, which in the case of the Probit model is a standard normal distribution function.

Then we proceed to estimate the wage penalties associated with WA and SOC(HE) measures through a modified version of the of the extended wage equation (Duncan and Hoffman, 1981) as proposed by Verdugo and Verdugo (1989). In this model educational mismatches are measured through dummy variables:

(3) $$lnw_i = \delta_o D^o{}_i + \delta_u D^u{}_i + \delta X'_i + \varepsilon_i$$

Where $D^o{}_i$ is the dummy variable for overeducation, $D^u{}_i$ the dummy variable for undereducation and X'_i is the row vector of controls including experience, experience squared, gender, working area, field of study and industry and other variables.

Although highly criticized, this specification has gained some popularity over the last two decades due to its capacity to describe the individual distribution of earnings among graduates' entering the labour market (Allen and Van der Velden, 2001; Dolton and Silles, 2008; Green et al., 2002; Sloane et al. 1999; Sloane, 2003; McGuinness, 2006; Green and McIntosh, 2007; Green and Zhu, 2010).

However, as we only refer to graduates, workers can be only overeducated. Thus, after dropping the undereducation term, our baseline specification is given by the following equation:

(4) $$lnw_i = \delta_o D^o{}_i + \delta X'{}_i + \varepsilon_i$$

Where $D^o{}_i$ is a dummy variable defined through SOC(HE)-Italy, assuming that an individual is overeducated if he/she is employed in job titles with a EQF Level that is undergraduate or lower (6 or below). Alternative specifications will be obtained by adding controls for industries, field of study and individual characteristics. In the first specification we include experience, experience squared, gender, working area, tenure, field of study and industry. In the second specification we add proxies of individual abilities related variables. In the third and last specification, social and family background proxies are included.

This estimate is affected by the selection bias that could raise from to the presence of inactive individuals in the sample. One of the standard estimation procedures for treating this selection problem is the two-step method proposed by Heckman (1979). However, in absence of appropriate exclusion restrictions, a multicollinearity problem is likely to arise (Puhani, 2000). Since the variables available in our dataset cannot fully address this issue we decide to use a subsample OLS estimate. Then we control for the selection bias by using propensity score matching technique (PSM) where the status of overeducated is the treatment variable. After checking that the common support condition is satisfied across more than 95% of treatment and comparison groups, we create matched "treatment" and "control" samples being identical in every other observable respect[12]. As some treated workers have lots of close neighbours and others only have one, we decide to use the kernel matching procedure in view of its capability to maximize precision without worsening bias. Finally, as a further robustness check, we re-estimate all the specifications with a series of Heckit regressions identified by three exclusion restrictions that have already been used in the empirical literature (Nicaise, 2001), although their exogeneity in the wage equation is questionable: living area, post-graduate studies or traineeships, and motherhood status[13].

6. Dataset and variables

The empirical analysis presented in this essay is based on data from the AlmaLaurea dataset on Italian graduates. The AlmaLaurea consortium collects every year extensive data on the graduates of each cohort and on their early working career path. This complex information is gathered in two stages. At the time of graduation students fill in a questionnaire providing their personal data, social and family background, educational path and performance, intrinsic motivation and other subjective features. Then, graduates are interviewed after one, three and five years from graduation and requested to provide information about their post-graduate education and training, and

12 Results of the distribution of propensity scores across treatment and comparison groups are available on request.

[13] Results of these estimations are available upon request.

in particular, about their current occupation. The latter information includes both objective and subjective items.

In our analysis we refer to the last cohort of graduates whose information is fully available for both steps of the survey. This cohort includes individuals having received either a two-year master degree or a five/six-year university degree (such as Medicine and Law faculties) during 2007, who completed their two step survey in 2012. The relevant population is composed by 184.669 graduates in 46 Italian universities, representing 61.5% of the Italian graduates in that year (source: National Institute of Statistics-ISTAT). The subsample of graduates who answered to the questionnaire after five year from graduation is composed by 31,162 individuals. Since we are only interested to individuals who are employed at the time of the survey, we exclude all graduates that declare to be either unemployed or inactive. Accordingly, we end up with a final sample of 25,523 graduates that declare to be employed at the time of the interview. However, since our results are limited to 6,219 individuals due e to missing data, our descriptive statistics are restricted to the same sample s.

Our main variables of interest are represented by the wage levels and by two dummies proxying the status of overeducated. Wages are measured in terms of net monthly earning. The graduate is requested to indicate his net earnings on a € 250 interval scale provided by the interviewer and ranging from € 250 to € 3000. Wage is then calculated as the mean of the intervals used in the questionnaire. Our measures of overeducation are based onto two items of the AlmaLaurea questionnaire. The SOC(HE) measure for overeducation is built on the occupational code, provided at a 5 digit level. The individual is considered as matched if his/her job is considered as a ‘graduate job’ according to the newly introduced SOC(HE)-Italy classification, as overeducated otherwise. The WA measure is based on a specific question for job requirements as by respondents. The individual is considered as matched if she/ he perceives her/his degree as either fundamental or useful for that particular job, as overeducated otherwise. The 5-digit occupational information, however, is only available in the interview realized after five year from graduation. Consequently, our empirical analysis is cross-sectional and refers to the survey carried out five year after graduation in 2012.

Additional variables included in the analysis include standard covariates of the human capital model: personal characteristics, educational path and achievements (field of study, graduation mark, and delay in completing the degree), employment history (experience, tenure). Individual heterogeneity is also captured by data on skills concerning software usage, foreign languages and the attainment of a scholarship, which are used as proxies of intrinsic abilities. In addition, we include variables related to current job's characteristics, such as the industrial sector, the working region, and the type of contract.

7. Descriptive statistics

Table 1 (see Statistical Annex) shows the overeducation incidence for the two measures. In our restricted sample, 23.2% of graduates are currently employed in jobs that require an undergraduate educational attainment according to the SOC(HE)-Italy classification, while 48.9%% of them declare to be overeducated in their current job (WA measure). Both percentages are higher when compared to the share of overeducated workers five years after graduation reported by most recent studies (Caroleo and Pastore, 2012; Ferrante et al., 2010; Verhaest and Van der Velden, 2013) for Italy and other European countries, in particular for the WA measure. Our different results can be explained by a large set of reasons: the deep economic crisis that hit Italy in 2011 and 2012, the entrance of new universities in the AlmaLaurea consortium in the same years, the different questions used to proxy overeducation by Caroleo and Pastore (2012) when using the AlmaLaurea questionnaire, the restricted sample we have taken into account for our empirical analysis (Table 1).

Applying the two measures of overeducation, two different groups of overeducated workers can be pointed out. The composition of these two groups of overeducated workers is quite different. On the one side, 71.5% of those overeducated according to SOC(HE)-Italy also perceive themselves as such. On the other side, only 42.1% of individuals stating to be overeducated are considered as such by the analysis on occupational levels. Hence, there is a low correspondence between the two measurements. This heterogeneity makes it meaningful to compare these indicators and discuss their validity on the ground that the two indicators can be actually rooted in different concepts. However it is important to emphasise that the pairwise correlation between the two measures of overeducation is medium sized and statistically significant (Table 2 in Statistical Annex). Notably, WA measure proves to be more strongly (and negatively) correlated with job satisfaction (-0.63) than the SOC(HE)-Italy measure (-0.49). This result supports our criticism that the WA measure is actually a proxy of job satisfaction rather than a valid measure of overeducation. Accordingly, we expect not only that individuals' ability and socio-economic background are among the main determinants of the WA measure, but also that the worker's view of the job post is highly related to his/her self-assessment, while this is not to be the case for the SOC(HE) (Table 2).

When considering individual characteristics we notice that gender acts in different ways on the basis of the different definitions of job requirements. The share of women employed in jobs that do not require a degree attainment (24.0%) is higher than the correspondent share of men (21.8%). Conversely, the share of men that perceive themselves as overeducated (53.5%) is higher than the correspondent share of women (46.0%). On the other hand, a favourable social and family background is associated with lower shares of overeducated workers under both definitions.

If one observes the field of study, the highest shares of matched graduates according to the SOC(HE) classification are in the fields of sciences, medicine and education, all of them showing poor overeducation incidences. At the opposite, the fields of economics and statistics, sport sciences, geo-biological, agriculture, and architecture show the highest shares of overeducated workers as defined by SOC(HE). These results partly differ from those associated with the WA measurements, which are higher for engineering and political sciences, and lower for agriculture and architecture.

When considering job characteristics, we focus on the contractual basis and the working area. In terms of geographical distribution, all Italian macro-regions show similar shares of overeducated workers (slightly more than 20%) without substantial differences between the two different definitions. Foreign countries are the only working area reporting a remarkably lower share of overeducated (15.4%). Consistently, the highest mean wage is reported by graduates working abroad. As far as the contractual basis is concerned, summary statistics show that self-employed workers are less overeducated than other employees in terms of both SOC(HE)-Italy and WA measures. Nonetheless, results for other types of contracts are ambiguous. Under both definitions, fixed-term and non-standard contracts are associated with lower levels of overeducation compared to open-ended contracts. However, the highest shares of overeducated workers (more than 50%) are associated with apprenticeships and temporary contracts.

Finally, overeducation incidence varies greatly across industries. We can describe the different sectors as grouped into three clusters. The first one, showing low levels of overeducation (less than 20%), includes education, information technology, health, professional consultancy, and other services. The second one, showing average levels of overeducation (between 20% and 30%), includes press, constructions, public administration and ICT. The last cluster, reporting high levels of overeducation (more than 30%), includes agriculture, energy and mining, manufacturing of metal products and machineries, wholesale and retail trade, transports and logistics, financial and insurance activities, arts, entertainment and recreation sector, other manufacturing. However, three of these "overeducated" sectors (manufacturing of metal products, other manufacturing, financial and insurance activities) are also characterized by wages that, on average, are higher than 1500 €. This evidence confirms the relevance of the industry sector as an expected determinant of both wages and overeducation no matter whether one adopts either the WA or the SOC(HE) measure.

8. Results

8.1 Determinants of overeducation and criterion validity of SOC(HE)

In order to assess whether WA and SOC(HE) are convergent or not, we present results from a set of Probit estimates in which the dependent variable is the probability to be overeducated and the covariates are given by a set of explanatory variables along with an array of control variables. Through standard techniques of estimation, we calculate the marginal effects associated with each covariate (i.e.: the effect of an increase in the level of the covariate on the probability to be overeducated when all the other covariates assume their average value). By comparing the results of the different estimates we identify the effect of each covariate on the probability to be overeducated for both overeducation measurements. Before doing so, it is important to emphasize that when comparing the results of two alternative estimations, one has to take into account the sign and the statistical significance of the parameters as well as their magnitude. As pointed out by McCloskey and Ziliak (2004) size matters as it can be conceived as a crucial component of the economic significance of a parameter. Thus, when comparing two alternative parameters there's no reason to mark down just their signs and whether they are significant or not. It is crucial to analyze differences in magnitude too since this is the main indicator of different economic meanings.

Results from the two Probit models (Table 4 in Statistical Annex) can be summarized by identifying five types of covariates: two groups of background variables (macro and meso variables) such as 1) industry variables and 2) geographical variables; and three groups of micro variables referring to labour demand and/or supply such as 3) job post characteristics variables; 4) field of study variables, which can be interpreted as both demand and supply indicators, and 5) variables about graduates' individual characteristics and performances during Higher Education (HE) studies. The differences between the five types of variables are quite striking:

1) the marginal effects of industry dummies are statistically significant for most cases and show equal signs. However, differences in magnitude are in a few cases extremely high;

2) the diversity of the marginal effects of geographical variables is very pronounced in terms of both the statistical significance of the parameters and their magnitude;
3) job characteristics variables are all significant at least at 10% when the SOC(HE) measure is adopted, whereas for WA both the fixed-term contracts (FIX_CON) and the non-standard contracts (NONST_CON) are not significant. Interestingly, tenure (which is highly correlated with wages and high levels of job satisfaction) is statistically significant for WA but not for SOC(HE);
4) remarkable differences can be observed for the field of study both in terms of statistical significance and of magnitude. Surprisingly, with only two exceptions the magnitude of parameters calculated for SOC(HE) is higher than that observed for WA;
5) some striking differences can be observed as far as the individual characteristics are concerned. What happens for the gender dummy is quite remarkable. This variable is negative and highly significant for SOC(HE) and positive while it is not significant for WA. Another noteworthy diversity can be detected for the variable indicating that at least one parent has a university degree, which is negative for SOC(HE) and not statistically different from zero for WA.
All in all, the deep discussion of the Probit models points out how same variables differently affect the two alternative measures of overeducation. Basically, this provides evidence that the variability of the two indicators is affected differently by the same variables. In this way the analysis of the determinant of the two measurements of overeducation through the Probit models result into a process of discriminant validation emphasizing the diversity of the two indicators[14]. In our view, this is due to the fact that

[14] In order to test the robustness of the Probit specifications, we have also run a bivariate Probit model, in which the determinants of both indicators of overeducation are estimated jointly. As the parameters in the overeducation equation in the univariate and bivariate probit estimates are almost identical in all specifications, we have decided to present and discuss the results for the univariate probit only.

subjective measures of overeducation can reflect individual perceptions and expectations as shown by their higher correlation with job satisfaction proxies reported in paragraph §4. This is confirmed by the estimates of parameters related to a variable indicating the degree of job satisfaction. This variable is negative and highly significant for WA and negative, but only at 10% significance, for SOC(HE) as reported in the fourth column of table 4. Besides, the parameter for SOC(HE) is smaller, in absolute value, than the one estimated for WA. Consistently, individual characteristics such as gender play a very different role as determinants of overeducation on the basis of the measure adopted. This huge difference can be explained with different expectations in terms of returns to education between men and women. On the other hand, the higher magnitude of the parameters attached to industries and fields of study for the SOC(HE) can reflect the fact that this measure is computed at the occupational level (see Table 3 in Statistical Annex).

8.2 Wage penalties and construct validity

When assessing wage penalties associated with WA and SOC(HE) measures (Tables 4) we compare three different specifications for each measure, following basically the same steps adopted for the Probit estimates. Whatever the measure adopted, overeducation has a negative and statistically significant impact on wages in all three specifications, as expected. Overeducated graduate workers earn, ceteris paribus, lower wages than their matched counterparts. Namely, the wage penalty ranges from 7.8% for SOC(HE) overeducated workers to 8.0% for WA overeducated workers, meaning that the WA parameter is slightly higher than the SOC(HE) parameter. Both estimates are consistent with part of the pre-existing literature (Allen and Van der Velden, 2001; Di

Pietro and Urwin, 2006) and in line with international comparisons showing that Italian overeducated graduates suffer one of the lowest pay-penalty in Europe (Ferrante et al., 2010; Barcena et al., 2011). This reduced gap can be explained by the poor performanceof wages in the Italian graduate labour market in the last decade as it is confirmed by large wage premia earned by Italian graduates working abroad in the AL sample (+45%).

Additionally, the gap between the two measures changes across the different specifications. When including ability proxies, namely high school mark (HSCH_MARK) and computer skills COMP_SKILL), we find that wage penalties decrease. This means that omitting observed skill heterogeneity (as done in the first specification) results in an upward bias of the impact on wages of both measures of overeducation. This is especially true for the WA measure as its wage penalty drops slightly more than that of SOC(HE) across the different specifications. Therefore SOC(HE) seems to be more stable than WA when trying to address the unobserved ability issue. Our ability proxies can capture only a fraction of the overall effect attached to skill heterogeneity as we can see from the relatively low change in overeducation coefficients and the non-significance of their effects on wages. Still, this indicates that one can expect a substantial divergence between the two measures and a lower correlation between individual ability and the SOC(HE) measure of educational mismatch when the omitted ability bias is fully taken into account. On the other hand, when including proxies of graduates' socio-economic background (third specification) we observe that SOC(HE)-Italy decreases significantly whereas the penalty associated with WA remains almost unaltered (Table 4).

PSM estimates substantially confirm these results[15]. The ATT coefficients are very similar to those presented before, both reporting a negative and significant effect of the treatment variable for both measures of overeducated. Such coefficients can be considered as estimates of the effect of being overeducated on wages. These estimates also confirm the existence of a slight difference between WA and SOC(HE) measures. Notably, the size of the penalty suffered by WA overeducated is still slightly higher than the one attached to the SOC(HE) measure. These statistical results are confirmed when we repeat the estimates using the nearest neighbour procedure. Overall, our treatment and comparisons appear to be rather similar after the matching, with no significant statistical differences in the means of the reported values[16].

In conclusion, our estimates of the earning function using two different indicators of overeducation brings two relevant results.

First, although the negative effect of both measures on wages is consistent with our expectations, the estimates of the direct effects of overeducation on the wage penalty are slightly different. Namely, the wage penalty associated with the WA measure is higher than that of SOC(HE) in all the specifications of the model. However, the gap is quite small especially if compared to the high divergence in the incidence and the determinants of the two measures. This may depend on two different factors. Firstly, even though the correlation between the two measures is not large, it is medium-sized and statistically significant. Secondly, this narrow gap can result from an omitted variable bias problem together with the inclusion of a wide set of controls concerning individuals' educational attainments and socio-economic background, abilities proxies,

[15] Although the balancing for individual covariates across the two groups is not always fulfilled, the imbalance for the covariate is always limited to one block of propensity scores out of five.
[16] To assess the quality of the matching we estimated the differences between the mean values of a subset of the covariates which are used to match the treatment and control groups. Detailed results are available on request.

and job characteristics. In fact, in the pre-existing literature on overeducation there is copious evidence suggesting that unobserved individual heterogeneity, such as unobservable abilities, can lead to an omitted variable bias problem. As McGuinness (2006) points out:

> "It is obvious from the results that the more poorly specified the model the more upwardly biased the overeducation penalty will be. The results demonstrates the importance of including job characteristics and some form of ability heterogeneity control"

The findings of McGuinnes (2006) are consistent with the results we have discussed a few lines above: the wage penalty associated with overeducation decreases no matter what definition is adopted when additional controls are added in the second and third specification of the model. However, penalties associated with WA and SOC(HE)-Italy decrease at different rates depending on what controls are included. Namely, as noted above, WA decreases more than SOC(HE)-Italy between the first and second specification in which ability proxies are included. Thus, unobserved individual ability could alter the effect of WA on wages in a more severe way than SOC(HE) due to the higher correlation of the subjective measure of overeducation with the omitted variables. This outcome suggests that the omitted variable distortion could be strong enough to widen the differences between the two measures of overeducation and further research should be carried out in order to assess whether the small gap between the two in terms of wage penalty remains constant when accounting for unobservable individual heterogeneity.

Unfortunately, the unavailability of longitudinal data concerning job descriptions and overeducation in the AlmaLaurea dataset, does not allow us to take into account unobserved individual heterogeneity and to fully address the omitted variable bias. Anyway, while waiting for further research in this direction it is important to stress that even at this stage the wage penalty is influenced differently by the same controls supporting the idea that WA and SOC(HE)-Italy capture different things. Namely, WA measures of overeducation are influenced to a greater extent by individual characteristics. This can be due to the fact that the individual assessment accounts for individual perception, expectations and, possibly, misperception. As expected, the individual characteristics and the educational career play a more relevant role in the analysis of overeducation as defined by WA than in that of SOC(HE). These factors, together with unobserved skill heterogeneity, affect the way in which graduates perceive their job position or their relative position in the labour market, and thus their perception to be overeducated. We thus consider WA to be more than a simple indicator of educational mismatch as it accounts also for graduates' expectations and perception of their relations to either the job or the labour market. Conversely, the wage penalty associated with SOC(HE) overeducation seems to be rather related to job characteristics and to the employment status of the worker rather than to individual characteristics while it is more sensitive to graduates' socio-economic background. Additionally, substituting one indicator for the other changes the effect of other variables on wage penalty, indicating that the two distinct indicators interact differently with some control variables.

Other differences between the two measures arise when we estimate the earning function within each of the major occupational groups. Omitting the categories of services support and manual occupations, for which SOC(HE) measure does not admit any matching with the degree attainment, we find that the wage penalty attached to the

WA measure is subject to wider changes across different occupations. Moreover, if we disaggregate the sample by occupational status the wage effect is significant only in the case of worker’s self-assessment. All in all, the divergent effect of control variables and the heterogeneity of the wage penalty across occupational groups supports the view that the two measures cannot be considered as equivalent when determining the wage effects of overeducation among graduate workers.

9. Conclusions

This paper shows the effects of microeconomic factors on the individual worker's likelihood to be overeducated within a limited time span. Secondly, the paper develops the analysis of the negative effect of overeducation on graduates' wage, taking into account the sample selection bias arising from the use of the wage as dependent variable. The analysis has been conducted by using two different measurements of overeducation. The first one (WA) is a standard worker's self-assessment measure. The second one (SOC(HE)-Italy) is based on a mixed strategy of measurement, which combines workers' self-assessment and job analysis, as described in details in § 2 and in Cattani et al. (2014). Results show that the two measures identify two substantially different sub-groups of overeducated workers. As a matter of fact, when we compare the two measures, WA overeducated workers are twice the number and less satisfied than their SOC(HE) counterparts. Furthermore, when analysing the determinants of overeducation, we find evidence that the effects of some explanatory variables partially change according to the measure of overeducation adopted. Therefore, as the variability of the two indicators is, albeit moderately, affected differently by the same variables (divergent validation) we can claim that the different approaches to the empirical analysis of overeducation give rise to different measures and point to different underlying phenomena. This is consistent with the medium-sized degree of linear correlation between the two measures.

However, the estimates of the wage penalty equation show that the magnitude of the wage effect is quite similar: both measures are associated with a statistical significant penalty that ranges between 7.5% and 8.5%.

Comparisons between these two kinds of results should be carried out cautiously. Nevertheless, both the analysis of the determinants of overeducation and the estimates of the wage equations suggest that the WA variable is driven by the individual characteristics of the employees and their position in the workplace. On the one hand, the high incidence of WA overeducation among Italian graduates is associated with the high share of individuals that are not satisfied with their job. This happens because self-assessment accounts for individual perception and expectations and can lead to biases due to misperception. Notably, individuals could perceive their job as inadequate to their educational level, in fact, basing their evaluation on poor graduate wage premiums even if the cognitive content of the assigned tasks is in line with their studies. On the other hand, measures of overeducation based on workers' self-assessment are influenced, as noted above, to a greater extent by individual characteristics.

On the contrary, in the SOC(HE) indicator, which is a combination of worker's self-assessment and job analysis, the subjective component is minimised and not correlated with the expectations and perceptions of the individuals, as the workers interviewed for the construction of this indicator do not belong to the sample of graduates interviewed. This confers the SOC(HE) measure a substantial advantage over the WA indicator in terms of construct validity. For the same reason, SOC(HE) might as well minimise the inevitable problems of endogeneity arising from reverse causality between the level of wages and a subjective indicator of overeducation. This is especially important in Italy where the pre-existing studies on these topics are only based onto WA measures and thus onto a possible misperception related to individual characteristics. The validity of the SOC(HE) measure also requires, however, a regular update of the classifications scheme in order to take into account the evolution of the cognitive content of each job title.

Overall, some of the results corroborate earlier findings in the literature, whereas others diverge. Like a typical JA measure, SOC(HE) avoids the correlation between over-education and workers' outcome variables, such as job satisfaction and show a lower incidence of overeducation. Together with the analysis of the determinants of overeducation, this result suggests that this phenomenon is less ubiquitous than the use of WA measures indicates and that so far most of the literature has probably assessed an occupational misperception rather than an educational mismatch. Conversely, the slightly higher correlation between SOC(HE) and WA, compared to the previous comparisons between subjective and objective measures, can be explained by the SOC(HE) reliance on workers' assessment rather than on job experts' classification. On the other hand, despite the divergence between the two measures, the educational mismatch is still associated with a wage penalty even if we measure it by taking into account the cognitive content associated with the constituent tasks of the job posts that graduates are currently holding. The evidence that both measures bring similar effects on the level of wages is in line with the literature and can be referred to the omitted variable bias that plagues many empirical contributions on this topic due to the relevance of individual heterogeneity and not to specific drawbacks in the validity of the SOC(HE) measure. Accordingly, a longitudinal study capable to address the omitted variable bias and based onto our mixed measure would greatly benefit the state of our knowledge regarding graduates careers and wages by assessing the extent to which the educational mismatch brings a wage penalty and the extent to which poor wages bring to occupational misperceptions. Finally, the interaction of WA and SOC(HE) measures with the other variables of the earning function is mostly divergent. Our new measure thus better acknowledges the importance of the demand side of the labour market thanks both to the construction procedure of the measure and to the wide set of information provided by the AlmaLaurea dataset. This also indicates that SOC(HE) classification is actually suitable to reflect labour market characteristics and skills' requirements that were in place at the time of the survey on Italian graduates.

References

Adcock R., Collier D. (2001), Measurement Validity: A Shared Standard for Qualitative and Quantitative Research, *American Political Science Review*, vol. 95, n. 3, pp. 529-546.

Allen, J.; Van derVelden, R. (2001) Educational Mismatches versus Skill Mismatches: Effects on Wages, Oxford Economic Papers n. 3 pp. 434-452.

Barcena-Martin E.; Budria S., Moro-Egido A. I. (2011), Skill mismatches and wages among European university graduates, MPRA Paper No. 33673.

Battu, H.; Belfield, C.; Sloane, P. (1999) Overeducation among graduates: A cohort view, *Education Economics*, vol. 7 issue 1, pp. 21–38.

Battu, H.; Sloane, P.; Building, E.; Street, D.; Park, S. (2004) Over-education and ethnic minorities in Britain, in *Manchester School*, vol. 72 issue 4, pp. 535–559.

Berg, I. (1970) *Education and Jobs: The Great Training Robbery*, New York, Praeger.

Caroleo F.E., Pastore F. (2012) Overeducation at a glance. Determinants and wage effects of the educational mismatch, looking at the AlmaLaurea data, Crisei Discussion Paper n. 18/2012.

Cattani, L.; Elias, P.; Purcell, K. (2014) SOC(HE)-Italy: a classification for graduate occupations, Working Paper DSE N° 963.

Cohen J. (1988), S*tatistical power analysis for the behavioral sciences*, second edition, Hillsdale NJ, Erlbaum.

Cohen J. (1992), A power primer, *Psychological Bulletin*, vol. 112, n. 1, pp. 155-159.

Di Pietro, G.; Urwin, P. (2006) Education and Skills mismatch in the Italian Graduate Labour Market, *Applied economics*, 38(1), 79-93

Dolton, P.; Silles, M. (2008) The effects of overeducation on earnings in the graduate labour market, *Economics of Education Review*, vol. 27 n. 2, pp. 125-139.

Duncan, G. and Hoffman, S. (1981) The Incidence and Wage Effects of Overeducation, in Economics of Education Review 1 (1), pp 75-86.

Elias, P., Purcell K. (2004). SOC(HE): a classification of occupations for studying the graduate labour market. Research Paper 6, University of Warwick, Institute for Employment Research.

Elias, P.,Purcell K. (2011) Higher education, intergenerational mobility and earnings: the case of the UK. Working Paper for ESRC.

Elias, P., Purcell, K. (2013) Classifying graduate occupation for the knowledge society, HECSU working paper n. 5/2013.

Frank, R. (1978). Why women earn less: the theory and estimation of differential overqualification, *American Economic Review*, vol. 68 issue 3, pp 360–373.

Ferrante F., McGuinness S.,. Sloane P. J (2010) Esiste «overeducation»? Un'analisi comparata, in Consorzio InterUniversitario AlmaLaurea (Eds.) XII Rapporto sulla condizione occupazionale dei laureati. Investimenti in capitale umano nel futuro di Italia ed Europa, Il Mulino, Bologna.

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

Gottschalk, P. and Hansen, M. (2003). 'Is the Proportion of College Workers in NonCollege Jobs Increasing? ', *Journal of Labor Economics*, Vol. 21, pp. 449-471.

Green, F.; McIntosh, S. (2007) Is There a Genuine Under-utilization of Skills Amongst the Overqualified?, *Applied Economics*, vol. 39 pp427- 439.

Green, F.; McIntosh, S.; Vignoles, A. (2002) The Utilization of Education and Skills: Evidence from Britain, *The Manchester School* vol. 70 n. 6, pp. 792-811.

Green, F.; Zhu, Y. (2010) Overqualification, Job Dissatisfaction and Increasing Dispersion in the Returns to Graduate Education, *Oxford Economic Papers*, vol. 62 pp 740- 763.

Groeneveld, S.; Hartog, J. (2004) Overeducation, wages and promotions within the firm, *Labour Economics*, 11(6) pp. 701-714.

Groot, W., Maassen van den Brink, H. (1997) Allocation and the Returns to Over-education in the UK, *Education Economics*, vol. 5 issue 2, pp. 169–183.

Groot W., Maassen van den Brink H. (2000), Overeducation in the labor market: a meta-analysis, *Economics of Education Review*, 19, pp. 149–158.

Halaby, C. (1994) Overeducation and skill mismatch, *Sociology of Education*, pp 47-59.

Hartog, J. (1980) Earnings and capability requirements, *Review of Economics and Statistics*, vol 62 issue 2, pp. 230 – 240.

Hartog, J. (2000) Overeducation and earnings: where are we, where should we go?, *Economics of Education Review*, Vol. 19 issue. 2, pp 131-147.

Hartog, J. and Tsang, M. (1987) Estimating, testing and applying a comparative advantage earnings function for the us 1969-1973-1977. Research Memorandum 8709, Universiteit van Amsterdam, Department of Economics.

Istat (2009) Indagine campionaria sulle professioni, Roma.

Kiker, B. F., Santos, M. C. (1991) Human capital and earnings in Portugal, *Economics of Education Review*, vol 10 issue 3, pp. 187 – 203.

Leuven, E. and Oosterbeek, H. (2011) Overeducation and Mismatch in the Labor Market, in Hanushek E.A., Welch F. (Eds.), *Handbook of the Economics of Education*, Vol. 4, Elsevier, Amsterdam.

Mason, G. (1995) The New Graduate Supply-shock. Recruitment and Utilisation of Graduates in British Industry, in Report Series Number 9, London, National Istitute of Economic and Social Research.

McCloskey D.N.and Ziliak S.T., (2004), Size matters: the standard error of regressions in the American Economic Review, *The Journal of Socio-Economics*, vol. 33, n. 5, pp. 527–546.

McGoldrick, K. and Robst, J. (1996) Gender differences in overeducation: A test of the theory of differential overqualification, *American Economic Review*, vol. 86 pp. 280–284.

McGuinness, S. (2006) Overeducation in the Labour Market, *Journal of Economic Surveys*, Vol. 20, No. 3, pp 387-418.

Nicaise I. (2001), Human capital, reservation wages and job competition : Heckman's lambda reinterpreted, *Applied Economics*, 33, pp. 309-315

Oosterbeek, H. and Webbink, D. (1996) Over scholing, overscholing en inkomen, *Economisch-Statistische Berichten*, vol. 81, pp. 240–241.

Purcell, L.; Elias, P. Atfield G., Behle H., Ellison R., Luchinskaya D., Snape J., Conaghan L., Tzanakou C. (2012) Futuretrack Stage 4: transitions into employment, further study and other outcomes, Warwick Institute of Employment Research, Mimeo.

Puhani P. A. (2000) The Heckman correction for sample selection and its critique, *Journal of Economic Surveys*, 14(1), 53-67.

Ramirez, A. A. (1993) Mismatch in the Spanish labor market: overeducation?, *Journal of Human Resources*, vol. 28, issue 2, pp. 259–278.

Rumberger, R. (1987) The impact of surplus schooling on productivity and earnings, *Journal of Human Resources*, vol. 22, issue 1, pp. 24–50.

Sicherman, N. (1991) "Overeducation" in the labor market, *Journal of Labor Economics*, vol. 9 issue 2, pp. 101.

Sloane, P. J.; Battu, H.; Seaman, P. T. (1999) Overeducation, Undereducation and the British Labour Market, *Applied Economics*, vol. 31, pp. 1437 – 1453.

Thurow, L.C.; Lucas, R. F. B. (1972) The American distribution of income: a structural problem. A study for the Joint Economic Committee, US Congress, Government Printing Office, Washington DC.

Van der Meer P.H. (2006), The validity of two education requirement measures, *Economics of Education Review*, 25, pp. 211–219

Verhaest D., Omey E. (2006a), Discriminating between alternative measures of over-education, *Applied Economics*, 38:18, pp. 2113-2120,

Verhaest D., Omey E. (2006b), The impact of overeducation and its measurement, *Social Indicators Research*, 77, pp. 419–448.

Verhaest D., Omey E., (2010), The determinants of overeducation: different measures, different outcomes?, *International Journal of Manpower*, Vol. 31, Issue 6, pp. 608 - 625

Verhaest D., Van der Velden R. (2013) Cross-country Differences in Graduate Overeducation, *European Sociological Review*, 29 (3), pp. 642–653

Verdugo, R.; Verdugo, N. (1989) The impact of surplus schooling on earnings: some additional findings, *Journal of Human Resources*, vol. 24, issue 4, pp. 629-643.

Statistical Annex

Table 1-Overeducation incidence

Overeducation incidence	N	**%**
SOC(HE) measure	6,219	
Matched	*4,776*	*76.80%*
Overeducated	*1,443*	*23.2%*
WA measure	6.219	
Matched	*3,183*	*51.1%*
Overeducated	*3,036*	*48.9%*

Table 2-Selected pairwise correlations (Pearson coefficients)

	SOC(HE)	WA	Job satisfaction
SOC(HE)	-		
WA	0.32***	-	
Job satisfaction (JOB_SAT)	-0.49**	-0.63***	-

**Significant at 5%
*** Significant at 1%

Table 3-Probit estimates with WA and SOC(HE) overeducation as dependent variables. Marginal effects

	(1) WA	(1) SOC(HE)	(2) WA	(2) SOC(HE)	(3) WA	(3) SOC(HE)	(4) WA	(4) SOC(HE)
IND_AGRIC	1.121***	1.535***	1.119***	1.495***	1.022***	1.496***	1.122***	1.499***
	(0.2023)	(0.1985)	(0.2050)	(0.2014)	(0.1790)	(0.2019)	(0.2052)	(0.2019)
IND_PRINT	0.786***	-0.00977	0.792***	0.00150	0.847***	0.00419	0.801***	0.00595
	(0.1982)	(0.2325)	(0.1989)	(0.2331)	(0.1854)	(0.2332)	(0.1990)	(0.2333)
IND_ENER	0.433***	0.916***	0.432***	0.911***	0.454***	0.921***	0.432***	0.923***
	(0.1436)	(0.1647)	(0.1450)	(0.1666)	(0.1245)	(0.1668)	(0.1451)	(0.1669)
IND_CHEM	0.772***	0.866***	0.789***	0.871***	0.769***	0.871***	0.786***	0.868***
	(0.1592)	(0.1725)	(0.1621)	(0.1752)	(0.1401)	(0.1753)	(0.1621)	(0.1754)
IND_MET	0.794***	0.992***	0.831***	1.012***	0.849***	1.014***	0.830***	1.015***
	(0.1223)	(0.1427)	(0.1244)	(0.1451)	(0.1033)	(0.1454)	(0.1245)	(0.1455)
IND_ELECT	0.257	0.638***	0.255	0.646***	0.322*	0.667***	0.270	0.672***
	(0.2116)	(0.2419)	(0.2120)	(0.2431)	(0.1857)	(0.2435)	(0.2129)	(0.2443)
IND_OTHM	0.848***	0.878***	0.847***	0.854***	0.906***	0.858***	0.837***	0.855***
	(0.1324)	(0.1468)	(0.1340)	(0.1489)	(0.1143)	(0.1493)	(0.1342)	(0.1493)
IND_CONST	0.386***	0.923***	0.407***	0.976***	0.351***	0.971***	0.409***	0.980***
	(0.1481)	(0.1698)	(0.1507)	(0.1722)	(0.1109)	(0.1725)	(0.1508)	(0.1724)
IND_TRADE	1.101***	1.169***	1.108***	1.167***	1.109***	1.168***	1.099***	1.165***
	(0.1094)	(0.1230)	(0.1108)	(0.1248)	(0.0941)	(0.1250)	(0.1109)	(0.1251)
IND_TRANS	0.941***	0.976***	0.942***	1.003***	1.046***	0.999***	0.943***	0.996***
	(0.1877)	(0.1885)	(0.1886)	(0.1906)	(0.1486)	(0.1910)	(0.1888)	(0.1911)
IND_COMM	0.884***	0.170	0.894***	0.157	0.928***	0.165	0.893***	0.163
	(0.1351)	(0.1552)	(0.1366)	(0.1573)	(0.1181)	(0.1574)	(0.1367)	(0.1574)
IND_FIN	0.685***	1.261***	0.674***	1.260***	0.812***	1.267***	0.672***	1.266***
	(0.1157)	(0.1341)	(0.1170)	(0.1357)	(0.0952)	(0.1360)	(0.1171)	(0.1360)
IND_CONS	-0.207*	-0.126	-0.215*	-0.0943	-0.180*	-0.0949	-0.219*	-0.0985
	(0.1132)	(0.1465)	(0.1149)	(0.1482)	(0.1010)	(0.1483)	(0.1149)	(0.1485)
IND_INFOR	1.205***	0.360**	1.225***	0.329*	1.196***	0.331**	1.233***	0.337**
	(0.1380)	(0.1646)	(0.1417)	(0.1685)	(0.1219)	(0.1687)	(0.1419)	(0.1687)
IND_BSERV	0.897***	0.905***	0.893***	0.881***	0.927***	0.885***	0.889***	0.882***
	(0.1343)	(0.1458)	(0.1369)	(0.1486)	(0.1151)	(0.1489)	(0.1370)	(0.1489)
IND_PUB	0.0306	0.198	0.0480	0.247*	0.246***	0.247*	0.0523	0.248*
	(0.1052)	(0.1285)	(0.1074)	(0.1312)	(0.0870)	(0.1314)	(0.1075)	(0.1314)
IND_EDU	-0.298***	-0.384***	-0.293***	-0.374***	-0.241***	-0.378***	-0.288***	-0.377***
	(0.1043)	(0.1337)	(0.1059)	(0.1357)	(0.0850)	(0.1359)	(0.1060)	(0.1359)
IND_HEAL	0.0618	0.184	0.0570	0.199	0.160*	0.207	0.0612	0.209
	(0.1095)	(0.1334)	(0.1109)	(0.1350)	(0.0898)	(0.1352)	(0.1110)	(0.1353)
IND_CULT	0.640***	0.526***	0.645***	0.542***	0.782***	0.534***	0.654***	0.539***
	(0.1438)	(0.1555)	(0.1448)	(0.1570)	(0.1123)	(0.1572)	(0.1450)	(0.1572)

	(1) WA	(1) SOC(HE)	(2) WA	(2) SOC(HE)	(3) WA	(3) SOC(HE)	(4) WA	(4) SOC(HE)
IND_OTHSER	0.281**	0.213	0.288**	0.235	0.349***	0.226	0.285**	0.224
	(0.1209)	(0.1456)	(0.1226)	(0.1476)	(0.0921)	(0.1479)	(0.1227)	(0.1479)
SELF_EMPL	-0.578***	-0.649***	-0.565***	-0.633***	-0.568***	-0.618***	-0.566***	-0.616***
	(0.0616)	(0.0719)	(0.0627)	(0.0729)	(0.0630)	(0.0732)	(0.0630)	(0.0732)
NOST_CON	-0.0540	0.483*	-0.0524	0.472*	-0.0598	0.466*	-0.0598	0.462*
	(0.2544)	(0.2567)	(0.2547)	(0.2566)	(0.2549)	(0.2563)	(0.2549)	(0.2564)
OTH_NOST	0.107	-0.355***	0.0890	-0.347**	0.0872	-0.326**	0.0872	-0.334**
	(0.0894)	(0.1377)	(0.0911)	(0.1393)	(0.0911)	(0.1399)	(0.0911)	(0.1401)
FIX_CON	-0.0191	-0.548**	-0.0105	-0.535**	-0.00597	-0.526**	-0.00122	-0.522**
	(0.2586)	(0.2626)	(0.2591)	(0.2628)	(0.2593)	(0.2625)	(0.2597)	(0.2626)
FULL_TIME	-0.190***	-0.202***	-0.202***	-0.213***	-0.204***	-0.212***	-0.195***	-0.205***
	(0.0585)	(0.0622)	(0.0596)	(0.0635)	(0.0597)	(0.0636)	(0.0598)	(0.0637)
TR_CONTR	0.280**	0.552***	0.290**	0.580***	0.293**	0.591***	0.283**	0.585***
	(0.1279)	(0.1236)	(0.1297)	(0.1253)	(0.1299)	(0.0131)	(0.1301)	(0.1258)
MALE	0.0330	-0.162***	0.0404	-0.155***	0.0439	-0.152***	0.0446	-0.152***
	(0.0421)	(0.0453)	(0.0433)	(0.0466)	(0.0434)	(0.0468)	(0.0434)	(0.0468)
AREA_NW	-0.0738	-0.0634	-0.0440	-0.0535	-0.0486	-0.0437	-0.0481	-0.0431
	(0.0544)	(0.0600)	(0.0556)	(0.0612)	(0.0557)	(0.0615)	(0.0558)	(0.0615)
AREA_SOU	0.211***	0.0463	0.222***	0.0340	0.221***	0.0265	0.221***	0.0277
	(0.0610)	(0.0682)	(0.0621)	(0.0694)	(0.0621)	(0.0695)	(0.0622)	(0.0695)
AREA_ABR	-0.159	-0.532***	-0.151	-0.516***	-0.155	-0.498***	-0.155	-0.498***
	(0.1030)	(0.1262)	(0.1088)	(0.1330)	(0.1090)	(0.1335)	(0.1091)	(0.1336)
POST_GRAD	-0.143***	-0.172***	-0.143***	-0.174***	-0.143***	-0.172***	-0.143***	-0.171***
	(0.0246)	(0.0316)	(0.0250)	(0.0322)	(0.0250)	(0.0321)	(0.0251)	(0.0321)
AGE	0.0851***	-0.0260	0.095***	-0.0184	0.095***	-0.0258	0.095***	-0.0260
	(0.0250)	(0.0288)	(0.0265)	(0.0308)	(0.0266)	(0.0309)	(0.0266)	(0.0310)
AGE_SQ	-0.001***	0.000381	-0.001***	0.000282	-0.001**	0.000352	-0.001***	0.000353
	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
EXP	0.678***	0.356*	0.741***	0.327	0.743***	0.324	0.730***	0.314
	(0.1825)	(0.2030)	(0.1859)	(0.2066)	(0.1860)	(0.2068)	(0.1861)	(0.2070)
EXP_SQ	-0.076***	-0.0432	-0.085***	-0.0385	-0.085***	-0.0383	-0.083***	-0.0368
	(0.0246)	(0.0272)	(0.0251)	(0.0276)	(0.0251)	(0.0277)	(0.0251)	(0.0277)
TEN	0.492***	-0.0362	0.498***	-0.0423	0.496***	-0.0470	0.495***	-0.0486
	(0.0450)	(0.0501)	(0.0458)	(0.0509)	(0.0459)	(0.0510)	(0.0459)	(0.0510)
FIE_AGRIC	-0.686***	0.174	-0.675***	0.242*	-0.677***	0.241*	-0.679***	0.241*
	(0.1224)	(0.1260)	(0.1256)	(0.1292)	(0.1257)	(0.1295)	(0.1258)	(0.1295)
FIE_ARCH	-0.298**	-0.217	-0.282**	-0.234	-0.281**	-0.235	-0.291**	-0.242
	(0.1359)	(0.1462)	(0.1403)	(0.1526)	(0.1402)	(0.1526)	(0.1403)	(0.1527)

	(1) WA	(1) SOC(HE)	(2) WA	(2) SOC(HE)	(3) WA	(3) SOC(HE)	(4) WA	(4) SOC(HE)
FIE_PHA	-2.248***	-1.389***	-2.186***	-1.272***	-2.182***	-1.263***	-2.180***	-1.261***
	(0.1324)	(0.1330)	(0.1387)	(0.1405)	(0.1388)	(0.1406)	(0.1389)	(0.1406)
FIE_ECO	-0.376***	-0.231***	-0.374***	-0.214***	-0.375***	-0.214***	-0.374***	-0.213***
	(0.0671)	(0.0718)	(0.0689)	(0.0733)	(0.0689)	(0.0734)	(0.0690)	(0.0734)
FIE_LAW	-0.487***	0.000507	-0.514***	-0.00923	-0.516***	-0.0110	-0.132	-0.00469
	(0.1070)	(0.1167)	(0.1088)	(0.1184)	(0.1088)	(0.1185)	(0.1587)	(0.1185)
FIE_ENG	-0.765***	-0.800***	-0.756***	-0.759***	-0.758***	-0.762***	-0.757***	-0.764***
	(0.0809)	(0.0925)	(0.0839)	(0.0959)	(0.0839)	(0.0961)	(0.0840)	(0.0962)
FIE_HUM	0.0598	0.0902	0.0208	0.0659	0.0215	0.0743	0.0245	0.0764
	(0.0752)	(0.0798)	(0.0765)	(0.0813)	(0.0765)	(0.0813)	(0.0766)	(0.0814)
FIE_MED	-0.318***	0.00675	-0.298***	0.0179	-0.297***	0.0129	-0.297***	-0.0125
	(0.0857)	(0.0963)	(0.0877)	(0.0982)	(0.0877)	(0.0983)	(0.0879)	(0.1049)
FIE_PSYCH	-0.380***	-0.316***	-0.400***	-0.328***	-0.405***	-0.332***	-0.400***	-0.329***
	(0.0697)	(0.0803))	(0.0712)	(0.0814)	(0.0713)	(0.0815)	(0.0714)	(0.0816)
PHD	0.430***	0.539***	0.439***	0.550***	0.443***	0.535***	0.445***	0.535***
	(0.1324)	(0.1547)	(0.1352)	(0.1575)	(0.1353)	(0.1578)	(0.1352)	(0.1580)
AV_MARK			0.146***	0.113**	0.147***	0.108**	0.151***	0.111**
			(0.0451)	(0.0493)	(0.0451)	(0.0493)	(0.0452)	(0.0494)
DEG_MARK			0.009**	0.015***	0.009**	0.014***	0.009**	0.0145***
			(0.0038)	(0.0044)	(0.0039)	(0.0044)	(0.0039)	(0.0044)
FATH_UPP					0.0459	-0.0696	0.0473	-0.0691
					(0.0395)	(0.0437)	(0.0395)	(0.0437)
PAR_DEG					-0.0569	-0.125**	-0.0558	-0.124**
					(0.0498)	(0.0556)	(0.0499)	(0.0556)
JOB_SAT							-0.010***	-0.0074*
							(0.0033)	(0.0039)
_cons	-3.035***	-1.205*	-4.481***	-2.413**	-4.476***	-2.149**	-4.424***	-2.088**
	(0.6086)	(0.7030)	(0.8342)	(0.9539)	(0.8367)	(0.9580)	(0.8374)	(0.9592)
Pseudo R^2	0.2330	0.2083	0.2357	0.2108	0.2360	0.2123	0.2371	0.2129
N	6219	6219	6065	6065	6065	6065	6065	6065

Reference variables for multiple dummy variables are the following ones: OTH_CONS for the industry, PERM_CON for the type of contracts, FIE_EDU for the field of study, AREA_CEN for the working area. Robust standard error in parenthesis

* Significant at 10%

**Significant at 5%

*** Significant at 1%

Table 4 - OLS estimates with natural logarithm of net monthly earning as dependent variable

	(1) lnW WA	(1) LnW SOC(HE)	(2) LnW WA	(2) LnW SOC(HE)	(3) LnW WA	(3) LnW SOC(HE)
OVERED	-0.0849^{***}		-0.0802***		-0.0804***	
	(0.0122)		(0.0123)		(0.0123)	
SOC(HE)		-0.0816^{***}		-0.0796^{***}		-0.0778^{***}
		(0.0127)		(0.0128)		(0.0128)
EXP	0.0292	0.0178	0.0101	-0.000211	0.0114	0.000843
	(0.0655)	(0.0658)	(0.0662)	(0.0664)	(0.0661)	(0.0663)
EXP SQUARED	0.00165	0.00285	0.00418	0.00526	0.00403	0.00514
	(0.0087)	(0.0088)	(0.0088)	(0.0088)	(0.0088)	(0.0088)
TEN	0.0763^{***}	0.0628^{***}	0.0731^{***}	0.0604^{***}	0.0731^{***}	0.0604^{***}
	(0.0128)	(0.0127)	(0.0129)	(0.0128)	(0.0129)	(0.0129)
MALE	0.173^{***}	0.169^{***}	0.170^{***}	0.166^{***}	0.169^{***}	0.165^{***}
	(0.0135)	(0.0135)	(0.0136)	(0.0136)	(0.0137)	(0.0137)
SELF_EMPL	0.802^{***}	0.820^{***}	0.805^{***}	0.822^{***}	0.798^{***}	0.816^{***}
	(0.1060)	(0.1064)	(0.1063)	(0.1067)	(0.1058)	(0.1062)
PERM_CONTR	0.920^{***}	0.934^{***}	0.921^{***}	0.935^{***}	0.922^{***}	0.935^{***}
	(0.1037)	(0.1043)	(0.1039)	(0.1044)	(0.1034)	(0.1039)
TR_CONTR	0.789^{***}	0.811^{***}	0.792^{***}	0.813^{***}	0.790^{***}	0.811^{***}
	(0.1063)	(0.1071)	(0.1064)	(0.1073)	(0.1059)	(0.1068)
NONST_CONT	0.643^{***}	0.670^{***}	0.646^{***}	0.672^{***}	0.647^{***}	0.673^{***}
	(0.1317)	(0.1326)	(0.1317)	(0.1325)	(0.1314)	(0.1323)
NONST_SELF	0.676^{***}	0.685^{***}	0.680^{***}	0.689^{***}	0.678^{***}	0.687^{***}
	(0.1076)	(0.1081)	(0.1079)	(0.1085)	(0.1074)	(0.1080)
OTH_NONSTAND	0.549^{***}	0.568^{***}	0.568^{***}	0.586^{***}	0.561^{***}	0.579^{***}
	(0.1198)	(0.1205)	(0.1197)	(0.1204)	(0.1192)	(0.1198)
FIX_CON	0.215^{***}	0.204^{**}	0.213^{**}	0.202^{**}	0.211^{**}	0.201^{**}
	(0.0831)	(0.0838)	(0.0830)	(0.0836)	(0.0832)	(0.0838)
REG_STUD	0.0688^{***}	-0.00301	-0.00339	-0.00297	-0.00292	-0.00252
	(0.0119)	(0.0065)	(0.0067)	(0.0067)	(0.0067)	(0.0067)
STUD_WORK	0.0316**	0.0650^{***}	0.0685^{***}	0.0649^{***}	0.0703^{***}	0.0666^{***}
	(0.0147)	(0.0119)	(0.0120)	(0.0120)	(0.0120)	(0.0120)
HSCH_MARK			-0.000361	-0.000325	-0.000340	-0.000304
			(0.0005)	(0.0005)	(0.0005)	(0.0005)
COMP_SKIL			0.00724	0.00770	0.00731	0.00778
			(0.0061)	(0.0061)	(0.0061)	(0.0061)
PAR_DEG					0.0132	0.0119
					(0.0156)	(0.0156)
FATH_ENTR					0.0272^{**}	0.0252^{**}
					(0.0118)	(0.0118)

	(1) lnW WA	(1) LnW SOC(HE)	(2) LnW WA	(2) LnW SOC(HE)	(3) LnW WA	(3) LnW SOC(HE)
UPP_CLASS					0.0108	0.00978
					(0.0177)	(0.0177)
_cons	5.933***	5.913***	5.951***	5.935***	5.938***	5.921***
	(0.1640)	(0.1639)	(0.1664)	(0.1663)	(0.1661)	(0.1660)
N	4921	4921	4791	4791	4791	4791
R2	0.3651	0.3635	0.3667	0.3655	0.3681	0.3667
Industries	Yes	Yes	Yes	Yes	Yes	Yes
Field of study	Yes	Yes	Yes	Yes	Yes	Yes
Type of high school	No	No	Yes	Yes	Yes	Yes
Working area	Yes	Yes	Yes	Yes	Yes	Yes

Reference variables for multiple dummy variables are the following ones: OTH_CONS for the industry, PERM_CON for the type of contracts, FIE_EDU for the field of study, AREA_CEN for the working area, HSCH_VOC for the type of high School. Robust standard error in parenthesis.
* Significant at 10%
**Significant at 5%
*** Significant at 1%