

# Determinants and wage effects of overeducation in Italy

## A comparison of five indicators of educational mismatch

**Emiliano Mandrone**

INAPP

**Francesco Pastore**

University of Campania Luigi Vanvitelli

**Claudio Quintano**

University of Naples Parthenope

**Debora Radicchia**

ANPAL

**Antonella Rocca**

University of Naples Parthenope

This paper aims to study dimensions of educational mismatch and quantify its effects on wages. Using the Inapp-PLUS survey, we measure both overeducation/skilling and undereducation/skilling in Italy, providing five different measures of the educational mismatch: three of subjective type and two of objective type. They are also synthesized in a single indicator, able to give a measure of the degree of severity of overeducation. These measures are provided both for university graduates and for upper secondary school graduates. Results highlight that the condition of overeducation is typical of males, younger workers, people coming from lower-income families, informal channels of recruitment and with a humanistic educational background.

*Il contributo analizza le dimensioni del disallineamento istruzione-lavoro e stima l'effetto sui salari. Utilizzando l'indagine Inapp-PLUS, misuriamo sia la sovra-educazione che il sotto-inquadramento in Italia, fornendo cinque diverse misure del disallineamento educativo: tre di tipo soggettivo e due di tipo oggettivo. Da questi si ricava un indicatore sintetico che misura l'intensità della sovra-istruzione. Queste misure sono stimate sia per i laureati che per i diplomati della scuola secondaria superiore. I risultati evidenziano che la condizione di sovra-istruzione è tipica degli uomini, dei lavoratori più giovani, delle persone provenienti da famiglie a basso reddito, che hanno usato canali informali di reclutamento e con un percorso scolastico umanistico.*

DOI: 10.53223/Sinappsi\_2022-03-7

### Citation

Mandrone E., Pastore F., Quintano C., Radicchia D., Rocca A. (2022), Determinants and wage effects of overeducation in Italy. A comparison of five indicators of educational mismatch, *Sinappsi*, XII, n.3, pp.130-155

### Keywords

Competences  
Overeducation  
Wage

### Parole chiave

Competenze  
Overeducation  
Salario

### Introduction

The under-utilisation of human capital also acknowledged as 'overeducation', is a growing phenomenon that is apparently at odds with the idea of a knowledge society. Even though the quantity of education (measured by educational qualifications and years spent in education) can be regarded as a partial indicator of the quality of human capital in terms of knowledge and skills, it seems paradoxical that the modest share of university graduates coming out of the Italian higher education system exceeds the still

scarce demand for qualified labour in the domestic market. In other words, the national economic system's need for graduates from higher education institutions is so modest that it fails in absorbing the yet scant number of graduates (compared to other OECD countries). This is in part the result of several decades of negative choices taken both by the private sector, with conservative business strategies scarcely inclined to innovation, and by the public sector, with an indiscriminate all-around distribution of financial resources. Among unfavourable conditions, then, an

impacting role is due to the lack of adequate investments in R&D, particularly over recent years. All these factors determined a waste of public and household resources and other negative effects such as the 'brain drain', low returns of education and scarce levels of technological innovation.

This paper aims to provide new measures and, to some extent, a new vision of the educational mismatch in Italy. The first element of novelty is in the newly available data from the last wave of Inapp-PLUS<sup>1</sup> (the Participation Labour Unemployment Survey of the National Institute for the Analysis of Public Policies), relative to 2014. This data bank allows us to measure both overeducation/overskilling and undereducation/underskilling, observable up to now only through objective (or statistical) methods. Overall, we provide five different measures of the educational mismatch, of which three are subjective and two are objective. We also introduce a new synthetic measure of mismatch, which allows us to partly overcome the limitations of the five above-mentioned in the direction of identifying the bulk or core share of individuals experiencing the mismatch.

Measurements will be provided not only for university graduates but also for upper secondary school graduates, where the latter measurements are rarely found in other Italian databases.

On this regard, it may be worth reminding that the Inapp-PLUS adopts the educational mismatch definition developed by Cedefop<sup>2</sup>, which allows us:

- i. comparing the rates of undereducation/underskilling across countries thanks to the adoption of the same definition in other EU countries and elsewhere;
- ii. reducing the bias of measuring the well-matched, which derives from ignoring undereducation/underskilling.

Moreover, Inapp-PLUS data provides detailed information on the educational paths and the school-to-work transition.

Besides providing new indicators of mismatch, through the econometric analysis we aim to find new evidence on the determinants of overeducation and

the measurement of the wage penalty associated with it using different estimation strategies. The OLS estimates provide the benchmark measures, but as well known in the literature, they are plagued by problems of possible omitted heterogeneity of the overeducated concerning the non-overeducated who are employed (endogeneity bias) and the non-employed who might also be overeducated if employed (sample selection bias). According to the relevant literature, the first problem of endogeneity is considered as not treatable with cross-section data, since there is no instrumental variable able to satisfy the two conditions of validity (especially the exclusion restriction) since wages and overeducation are two sides of the same coin – a given job-worker match – and, therefore, there is no variable which might affect the probability of overeducation without also affecting wages. Similar objections can be raised for sample selection bias because there are no instrumental variables able to affect the probability of finding a job, but not the wage provided by that job. We try to address both problems through a large number of instrumental variables thanks to the rich information on education and school-to-work transition provided by the Inapp-PLUS survey.

The outline of this paper is as follows. Section 1 provides a brief review of the literature dealing with the different ways in which the phenomenon has been measured so far. Section 2 contains an operational definition adopted in the individual queries of Inapp-PLUS and used to develop a 'synthetic indicator' for a new measurement of the mismatch between human capital and labour in Italy. Section 3 presents the descriptive statistics and comparisons between the five different mismatch indicators adopted and some typical determinants of overeducation. Section 4 presents the methodology proposed to study the determinants and wage effects of overeducation. Section 5 discusses the findings of estimates of determinants and effects. Some concluding remarks complete the paper.

## 1. Literature<sup>3</sup>

The following is a brief review of literature about main approaches adopted toward educational mismatch, both at theoretical and measurement level.

1 Former Isfol-PLUS, launched in 2005 by Isfol (Institute for the development of vocational training for workers): the survey's naming changed alongside Isfol's turning into Inapp (December 2016).

2 An overview of definitions about different types of educational mismatch can be found in Cedefop (2010).

3 For a full-length review on definitions, methodologies and estimates, see, among others, Hartog (2000), Quintini (2011) and Kucel (2011).

### **Theoretical approaches**

The analysis of educational mismatch has led to the development of a variety of theoretical approaches. A univocal definition of mismatch, which would be useful to conduct cross-country comparisons, has proven hard to identify. The main difficulties in the matter may be summarized in two points:

- i. the general approach adopted;
- ii. and the nature (subjective vs objective) of the indicators used.

The branches of research on the returns to human capital mainly look at the labour supply (Attanasio and Kaufmann 2017; Johnes 2019). The educational mismatch, acknowledged as the mismatch between the workers' actual qualifications and the level of education required to perform their current job, is considered the result of job seekers' behaviour: some individuals may choose to accumulate qualifications and develop their human capital to be the 'winners' in the search for a job (job competition model, Thurow 1979), thus ending up in a job position that does not necessarily require all of their investment in education to be used. Alternatively, other individuals may choose to accept a job, though being overqualified for it, in the hope of getting sooner a well-matched occupation and enhancing future career prospects and earnings while on the job. The latter is known as the career mobility theory and is an extension of the human capital model (Sicherman and Galor 1990). Job search theoretical models posit, instead, that individuals choose between waiting for a more appropriate job offer to come while remaining unemployed or, rather, taking a job for which they are overqualified if they feel they are less skilled<sup>4</sup>.

However, the human capital model remains the most important candidate for a theoretical explanation of the educational mismatch, although, in the early literature, erroneously, overeducation was depicted as a breach of the general validity of the human capital model. As, among others, Leuven and Oosterbeek (2011) point out, overeducation might still be a consequence of a lower-than-average human capital if despite increasing educational

attainment, the work-related component of human capital – general and job-specific competencies – is missing (see also Caroleo and Pastore 2018).

Many authors have investigated the main determinants of overeducation, and, among them, some have emphasized the role of the fields of study (Capsada-Munsech 2015; Montt 2017; Rossen *et al.* 2019).

Spatial differentials across regions are attracting the attention of observers. Davia *et al.* (2017), for instance, study the determinants of the educational mismatch across European regions. Moreover, interestingly, Croce and Ghignoni (2016) argue that lower labour mobility across regions might tend to make persistent geographical differentials in the educational mismatch.

Caroleo and Pastore (2018) suggested using the Heckman procedure as a device to disentangle each theoretical hypothesis from the others and have found that the human capital model and the job competition model are the most convincing explanations in the case of Italy.

### **Measurement issues**

The International Labour Office (ILO) has expressed the necessity of enhancing the efforts to converge to a definition and a taxonomy of mismatch that is recognized at the international level.

Vertical mismatch generically refers to a situation where the level of education obtained by a worker is higher (or lower) if compared to that required in the current job. Horizontal mismatch refers to a situation where the worker has the qualification level required, but different skills and knowledge than those required for the job currently performed<sup>5</sup>.

The Organisation for Economic Co-operation and Development (OECD) launched a specific 'New Skills for New Jobs' initiative aimed at the systematisation of definitions to obtain comparable measures and understand the skills mismatch in different countries, given the costs it generates and the need for designing policy measures to contain or minimize the phenomenon.

The most recent editions of the OECD 'Education at glance' (various years) report that Turkey is now doing better than Italy in terms of average levels of tertiary

4 For more detailed reviews on the theoretical and empirical literature, see Sloane (2003), Leuven and Oosterbeek (2011), Quintini (2011), Cappelli (2015), Caroleo and Pastore (2016).

5 Istat (2012) suggests some interesting insights and indications on the educational mismatch and its measurement.

education. Nevertheless, the level of overeducation/overskilling is more severe in Italy than in countries with higher education levels. This calls for a demand-side interpretation (Croce and Ghignoni 2012; Franzini and Raitano 2012) and, as pointed out in previous studies (Caroleo and Pastore 2013; 2018), suggests a difficulty of the school-to-work transition system in providing the competencies demanded by the productive system, due to insufficient interaction between the educational system and the world of work. On the demand side, the empirical results can only be explained by a still very traditional productive structure generating a low need for higher-level competencies. Moreover, due to its weak connections with the productive system, the educational system often fails to generate competencies immediately usable by firms.

A key element of the current debate on the matter rests in the importance of the ability to adjust competencies to adapt to the rapid changes occurring in the world of work. The educational mismatch has important economic implications at different levels: at the individual level, it affects the degree of job satisfaction and earnings; at the firm level, it reduces productivity and increases both turnover and the search for a different occupation; at the macroeconomic level, it alters the degree of tolerance to unemployment and reduces GDP growth through the loss of human capital and/or the fall in productivity. It is, therefore, important to identify the main determinants of the mismatch and to develop a comprehensive political strategy to limit such negative effects.

Skill mismatch between workers and jobs has also proven hard to measure. Only a few datasets contain detailed information on workers' skills and abilities and those required by their jobs. And even when containing such information, often just report on workers' levels of literacy and numeracy. The self-assessments on own skills are rare, too. As a consequence, most of the academic literature quoted in this paper focuses on the mismatch between obtained qualifications and those required in the current job. While education can be regarded as a good proxy for knowledge, there have been weak developments in the analysis of differences between qualifications and skills.

The second class of problems pertains to the measurement of the phenomenon. The generally used indicators can be divided into subjective and objective measures. The former category comprises measurements obtained from workers' responses to questionnaires (self-reported) on the requirements (direct assessment) and adequacy (indirect assessment) of their skills and qualifications in their actual jobs.

The weakness of such measurement strategies is in the possible biases introduced by the workers assessing their jobs, due to their expectations, involving their job satisfaction and in some cases an overestimation of their own technical and cultural knowledge. A correct balance between the two categories of measurement is functional to the distinction between actual overeducation (effective mismatch between qualification and job requirements) and spurious overeducation (perceived mismatch).

To control for the biases introduced by self-assessments, subjective measures are typically complemented with several objective indicators derived from the analysis of qualifications and years spent in education by a reference population in a given job (Franzini and Raitano 2012).

## 2. Indicators of educational mismatch

The Inapp-PLUS 2014 wave<sup>6</sup> (Mandrone 2008 and 2012; Giammatteo 2009; Corsetti and Mandrone 2012; Corsetti *et al.* 2014) provides additional, detailed, and comprehensive information on the phenomenon of overeducation, as defined by Cedefop, that suggests that overeducation is becoming ever more common and alarming. A strong point of the new definition is to make the Italian data comparable with those of other developed countries adopting the same definition. The choice of the dataset is also connected to the variety of controls it offers, which allows to study relations and links otherwise invisible to other data banks (Ortiz and Kucel 2008). Examples of information on local labour market efficiency that are studied from the dataset are the availability of jobs, local mobility, and commuting. This also allows for georeferencing the phenomenon under consideration (Buchel and Van Ham 2003).

<sup>6</sup> Inapp-PLUS is a survey on labour supply in Italy conducted on 55.000 individuals (CATI) with a share of panel interviews. Data are available at <https://inapp.org/it/dati/plus>.

**Table 1. Indicators of educational mismatch for upper secondary and university graduates**

Indicator	Definition		Upp. sec. school graduates	University graduates
<b>Over-under skill</b>	Query: "To what extent your professional skills correspond to those required in your current job? My professional skills are ..."	• Much higher	9.4	13.4
<b>Sheepskin effect</b>	Query: "Is your educational qualification necessary to access your current job?"	• Yes	62.7	79.3
		• No	37.3	20.7
<b>Over-under education</b>	Query: "In your opinion, which level of education is the most adequate to perform your current job?"	• Less than upp. sec.	16.3	2.2
		• Upper secondary	65.8	16.0
		• Bachelor degree	11.7	22.6
		• Master's degree	5.9	53.5
		• Graduate degree	0.4	5.7
		• Somehow higher	19.8	22.2
		• Roughly the same	67.7	62.1
		• Somehow lower	2.8	2.1
<b>ISCO based</b>	Coherence of the educational qualification of the i-th individual with the ISCO profession where s/he is classified.	• Under	20.5	-
		• Match	59.8	79.1
		• Over	19.7	20.9
<b>Mode based</b>	Coherence of years spent in education by the i-th individual with the modal value of the profession.	• > mode	19.8	36.5
		• = mode	80.2	63.5

Source: Authors' own elaboration on Inapp data

The Inapp-PLUS survey differs from usual surveys, mostly household-based, due to the absence of indirect respondents (proxy interviews) and for directly targeting the status perceived by respondents<sup>7</sup>. This approach made it possible to submit an extremely detailed questionnaire on the nature of employment, education and household conditions, allowing – through dedicated modules – the analysis of very specific phenomena. In this respect, the survey captures the actual situation of the labour market, as it is perceived by the individuals directly involved, with particular attention to the difficulties experienced by youth, women, and people over 50.

Table 1 offers a synoptic view of the educational mismatch indicators available in our data and a synthetic indicator. The structure is articulated around three concepts: Overskilling/Underskilling,

Credentialism, and Overeducation/Undereducation.

Overskilling/underskilling refers to the mismatch between skills and abilities held and those required in the job currently performed.

The overeducation status identified in row 2 is based on credentialism (defined also as the sheepskin effect in the Anglo-Saxon literature), namely the phenomenon by which we understand the tendency of employers to give a job to those who possess a given degree or educational qualification. When this happens, we consider the degree as a credential to access a job. More specifically, the sheepskin effect refers to the salary gap between those obtaining an educational qualification and those who, having spent the same number of years in education, did not obtain it (Hungerford and Solon 1987).

Overeducation/undereducation (row 3) refers to the mismatch between the obtained educational qualifica-

<sup>7</sup> The first consequence is that the characteristics of PLUS occupations are cleaned from the component of non-prevailing employed (*i.e.* those individuals that are considered as employed according to the official Istat-ILO definition – having worked for at least one hour over the last week – but with an economic activity not so substantial that they can be considered as employed 'tout court') and include some other categories of unemployed (or job-seekers) that, on the contrary, are not considered in the official Istat-ILO definition (potential workforce). However, it is always possible to retrace the labour force aggregates defined by Istat-Eurostat from PLUS data, by using *ad hoc* queries.



tion and that required in the job currently performed.

While these first three indicators are subjective ones (based on self-responses), the last two indicators are objective ones, derived from comparisons with other people employed in similar jobs (II-digit occupational classification). This approach takes into consideration both the aspects of job satisfaction and overestimation of competencies, as well as the limitations posed by rough classifications.

The measurement of those phenomena is provided for both upper secondary school (column 3) and university graduates (column 4) in Table 1. Some details on these indicators are provided below.

### ***The sheepskin effect***

The sheepskin effect in this context can be taken to measure the role played by an educational qualification to obtain a job in the labour market. A high share of the mismatch indicates the degree of competition for that particular job among people holding a given educational qualification. The higher the value of the mismatch, the higher the probability that an individual with a lower/higher educational qualification gets the job and, hence, the lower the specific added value generated by that educational qualification for those holding it. The effect concerns primarily a considerable share of upper secondary school graduates who consider their own educational qualification as a non-necessary requirement for accessing their current job: 1 upper secondary school graduate out of 3 (against 1 university graduate out of 5) answered that their job could be accessed also having a lower level of education. This is a clear symptom of a generic labour demand, which is just vaguely oriented towards high professionalism. The most striking aspect of this information is its high incidence among upper secondary school graduates, as compared to university graduates, even if also the latter is higher than that found in the literature (Silles 2008; Matkovic and Kogan 2012). Actually, in the case of upper secondary school graduates, there is no term of comparison, as the necessary statistical data are almost completely missing for most countries and certainly for Italy. Nevertheless, the Inapp-PLUS estimates are very high, even in absolute terms, and suggest that only a few jobs performed by upper secondary school graduates are exclusively accessible to them. This creates an ever-fiercer competition between

them and those with a lower/higher educational qualification. Upper secondary school diploma has become very common nowadays, although less than the EU average (OECD 2015), and is, consequently, perceived as not being very distant from compulsory education. This circumstance may also explain why, at present, many upper secondary school graduates tend to enrol for a higher qualification.

In the case of university graduates, the closest definition to the one we use is that of the AlmaLaurea dataset, although such a comparison is to be taken with due caveats, as the Inapp-PLUS data includes individuals of different ages, whereas the AlmaLaurea includes only individuals who received their degree recently. Using the information for university graduates in 2010, Caroleo and Pastore (2018) report a mismatch value of 16.5% after one year from graduation, roughly 13.0% after three years from graduation, and 11.4% after five years from graduation. Assuming that discrepancies with the AlmaLaurea data do not depend much on differences in definitions, but, rather, on the different ages of the interviewees, the likely conclusion is that, even for university graduates, the number of jobs they get and perform, though not requiring a university degree, is actually decreasing among the youngest generation, and the competitive threat posed by those having a different educational qualification is becoming lower with time passing. It is an indirect sign of an increase in the sheepskin effect.

### ***Overeducation / undereducation***

Following the Cedefop's definition, the second subjective indicator allows us to observe both over- and undereducation, by asking the surveyed persons to specify the level of educational qualification necessary to perform their current jobs. The definition adopted also allows us to obtain, for the first time in Italy, a measure of undereducation that is not based on the statistical method. This indicator measures the educational mismatch, suggesting the degree to which those holding a given educational qualification perform a job that is in line with, above or below their level of competencies. Among university graduates, overeducation seems to prevail (more than 18%) compared to undereducation (5.7%), again confirming the low demand for education and higher skills compared to the demand existing in the Italian productive system. On the contrary, for upper

secondary school graduates, the phenomenon of undereducation (roughly 18%) just exceeds that of overeducation (16.3%), further confirming the current devaluation of upper secondary educational qualifications. The breadth of undereducation among upper secondary school graduates suggests that many individuals holding such level of qualification still occupy higher job positions. The percentage figure relative to university graduates that we obtain is lower than the one reported by Cedefop (21.0%), which is obtained by applying the same definition of Inapp-PLUS to data from the European Union Labour Force Survey, as part of a 2012 study for the European Commission, covering the period 2001-2011 (Pouliakas 2012, 360, Table 7).

If this discrepancy was to be ascribed to economic factors, rather than to the different ways of data collection and the weights of the direct interviews in Inapp-PLUS, a reduction of overeducation for university graduates would emerge, even in presence of a sizeable increase in the share of university graduates in the last years up to 24% according to recent OECD estimates, in correspondence also with the persistence of the economic crisis started in 2007. Of course, such a result should be interpreted with caution, monitoring the evolution of this indicator in the years to come. Lastly, the low level of undereducation among university graduates confirms the fact that the available jobs rarely require levels of education higher than a bachelor's degree.

### ***Overskilling / underskilling (over\_skill)***

As regards the third subjective indicator, it is worth noting that underskilling presents a very low percentage both for upper secondary school graduates (3.2%) and for university graduates (2.3%), whilst the statistical measures tend to considerably overestimate the phenomenon, as we will soon show.

Something different happens for overskilling, which is oversized if compared with estimates from other studies: it reaches 35.6% among university graduates and 29.2% among upper secondary school graduates. McGuinness and Sloane (2010, Table 3.6) estimate 21% after one year from university graduation and 11% after five years from graduation. The strong overskilling registered in the Inapp-PLUS data may have some explanations. On the one hand, the demand for high-skill workers is quite low but,

on the other hand, the way teaching is organized in both schools and universities may provide a further explanation due to the very poor integration with the world of work. Caroleo and Pastore (2018), for instance, suppose that overskilling in Italy may be linked to the low spread of work-related competencies, among university graduates, because of the lack of work experience and the emphasis of university teaching on general education. Hence, this finding may also be a consequence of how the school-to-work transition is organized. The reduction of the figure in the AlmaLaurea data is probably due to the increasing skill requirement of the production system concerning the previous decades.

### ***ISCO based indicator***

The ISCO-based indicator is of an objective type. It measures the coherence of i-th individual's educational qualification by looking at the correspondence between each large professional group and a given qualification, as established by the International Standard Classification of Occupations, ISCO8. In this case, a mismatch may occur between the theoretical definition of the qualification required by a given profession and the qualification actually acquired by the surveyed person, independent of individual perception. As a matter of fact, overeducation is considered as a mismatch between the typical educational requirement of the profession, and hence a standard educational path, and the possible deviances, *i.e.* individuals with a non-standard educational path performing that profession – except for the first category (1 digit, the professional group 'entrepreneurs and executives') that is considered as being matched by definition. Being a statistical measure, it is not affected by perceptions. However, it may not necessarily correspond to a situation where the individual human capital is inadequate for the job performed. According to this indicator, one-fifth of the workers with a university or upper secondary school qualification is overeducated and 20% of workers with an upper secondary school qualification are undereducated. Such results signal a certain misalignment, both formal and substantial, between levels of education and qualifications necessary to perform a job. No undereducated university graduates are reported.

### Mode based

This indicator measures the matching between the level of educational qualification of the *i*-th individual and the modal value of the qualification per profession. The modal value becomes the benchmark qualification for that profession. Those workers who are in a different situation are considered as not well matched for the distribution of educational qualifications for that job. The last two objective indicators can be thought of as reporting mismatches from the typical or prevailing binomial profession-educational qualification. Also, in this case, the indicator does not necessarily signal job dissatisfaction. One-fifth of upper secondary school graduates and over a third of university graduates are overeducated, according to this indicator. No undereducation is found.

Compared to the corresponding subjective indicator, there is an overestimate for both categories of graduates, which is more severe in the case of university graduates. This overestimate may depend on the very method of calculation: one of the limitations of objective (or statistical) indicators is that they tend to return some overeducation anyway and that they are affected by the degree of statistical aggregation of considered groups, *i.e.* the higher the degree of aggregation of considered professions, the less representative is the modal value and the higher the share of those who are distant from the mode, in this way giving the impression of overestimating overeducation. It may be interesting to point out that, in the Italian case, both subjective and objective measures tend to return high values of the educational mismatch, showing the worrying diffusion of the phenomenon, especially in consideration of the relatively low level of educational attainment. However, the differences

in the extent of overeducation across different indicators are not extremely sizeable.

### Synthetic indicator and reclassification of overeducation

The presence, in the same sample survey, of five different indicators of the educational mismatch supplies an uncommon richness of opportunities for possible analyses. Note that in the following analysis we use the terms *overeducation lato sensu* as referring to the educational or occupational mismatch. Trying to overcome the limitations of both subjective and objective indicators, we propose a synthetic indicator constructed to maintain the best properties of each component indicator. The theoretical framework for mismatch classification is presented in Table 2.

The combination of intensity and type of mismatch identifies three conditions corresponding to 6 different levels:

1. *Matched*: namely employed individuals not presenting any mismatch in any of the 5 indicators used;
2. *Partly mismatched*: individuals identified as mismatched by 1-4 indicators;
3. *Mismatched*, individuals identified as mismatched by all the indicators used.

Figure 1 shows the levels of overeducation to which individuals may be exposed: “0” *well-matched* (perfect coherence between qualification and job); from “1 to 4” *partly mismatched*; and “5” *mismatched* (complete misalignment between qualification and job). The results hint at a very severe level of overeducation in a relatively small number of people (about 10%). This indicator

**Table 2. Types of mismatches**

Over-under skill	Sheepskin effect	Over-under Education	ISCO based	Mode based	Types
0	0	0	0	0	<i>Matched 0/5</i>
		1 factor out of 5 – 20%			<i>Partly Mismatched 1-4/5</i>
		2 factors out of 5 – 40%			
		3 factors out of 5 – 60%			
		4 factors out of 5 – 80%			
1	1	1	1	1	<i>Mismatched 5/5</i>

Source: Authors' own elaboration on Inapp data



envisions the coexistence of 5 factors out of 5 that, given the subjective and objective nature of the adopted indicators, should correspond to a 'highly probable' overeducation. This finding of a relatively low level of overeducation across different dimensions and indicators utters the findings of Flisi *et al.* (2017).

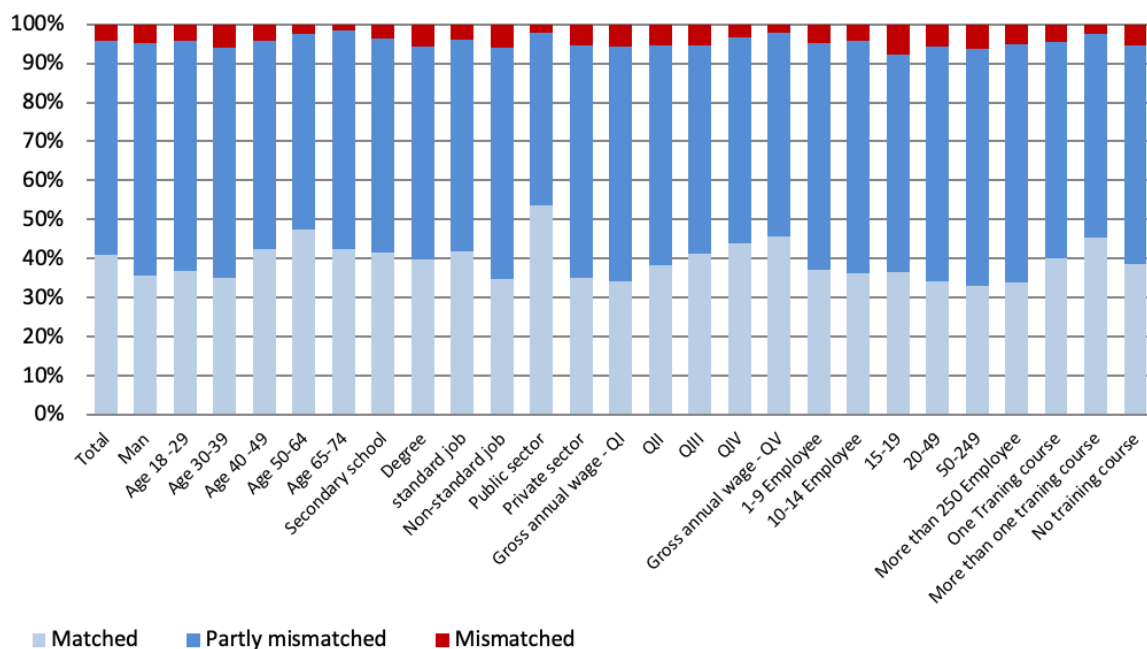
In more positive terms, we find better matching in the Public Administration, among women, individuals over 50, those with higher wages, and those engaged in training programmes (particularly in continuing vocational education and training, CVET).

Best matching results are observed for: women, who tend to choose jobs with higher security level and guaranteed paid maternity leave, then also offering better guarantees in terms of matching; adults, as prescribed by the career theory, since in the presence of overeducation, individuals move to better jobs if their qualifications are not recognized in present jobs or may benefit of career progression during their working life which determines a good match; public sector employees, as the public administration has to recognize the educational qualification both as an access condition and as a condition for the career development to senior positions.

### 3. Descriptive statistics

Similar to what is found in Figure 1, the phenomenon of overeducation, in its several aspects, shows significantly different levels by age class, educational qualification acquired, and gender (Table 3a and Table 3b). The educational mismatch decreases with age, suggesting that over time, even if slowly and partially, an adjustment between human capital and work eventually occurs, due to occupational mobility. In the age class 40-49, more than 43% of employed individuals report no factor of mismatch (35% in the age class 30-39), and the percentage figure increases to 48% for those over 50. The only element of mismatch that is difficult to be absorbed over time is provided by the overskilling indicator; in fact, we observe no significant percentage variations among different age classes. The gender analysis shows that mismatch is always more frequent among men for all the adopted indicators and over 47% of women employed report no factors of overeducation. The analysis of indicators of overeducation per educational qualification highlights that, roughly speaking, an upper secondary school graduate is generally less overqualified than a university graduate. Only the sheepskin effect involves a relevant share of upper secondary school graduates considering

Figure 1. The composition of overeducation for main controls



Source: Authors' own elaboration on Inapp data

**Table 3a. Overeducation for several controls, %**

	Tot	Age class				Gender		Education		
		18-29	30-39	40-49	50-64	Men	Women	High second.	Tertiary	
<b>1 Sheepskin effect</b>	<b>32</b>	39	39	31	24	36	28	37	21	
<b>2 Overeducation</b>	<b>17</b>	20	22	16	12	18	16	16	18	
<b>3 Overskill</b>	<b>31</b>	29	33	30	31	33	29	29	36	
<b>4 Over ISCO</b>	<b>20</b>	22	25	20	15	26	14	20	21	
<b>5 Over Mode</b>	<b>17</b>	21	23	16	11	19	14	20	37	
<b>Factors of overeducation (1-5)</b>	<b>0</b>	41	37	35	43	48	36	47	41	40
	<b>1</b>	26	24	25	25	27	25	26	25	26
	<b>2</b>	15	19	15	15	12	17	13	15	14
	<b>3</b>	9	11	12	8	7	11	7	10	8
	<b>4</b>	5	5	7	5	4	6	4	5	6
	<b>5</b>	4	4	6	4	3	5	4	4	6
		Employment Contract					Job satisfaction			
		Typical work	Atypical work	Indefinite contract	Temporary contract	Short term	Self employee	Collaborators	High satisfaction	Low satisfacti
<b>1 Sheepskin effect</b>		31	39	30	38	43	34	45	30	40
<b>2 Overeducation</b>		16	24	16	24	29	15	22	15	22
<b>3 Overskill</b>		31	34	31	32	38	31	36	29	37
<b>4 Over ISCO</b>		20	22	20	23	25	18	16	19	24
<b>5 Over Mode</b>		16	20	17	20	20	15	18	16	19
<b>Factors of overeducation (1-5)</b>	<b>0</b>	42	35	43	37	31	40	31	43	34
	<b>1</b>	26	25	25	24	22	27	30	26	24
	<b>2</b>	14	17	14	16	16	16	19	14	17
	<b>3</b>	9	11	9	10	19	10	10	9	11
	<b>4</b>	5	6	5	6	8	4	6	5	7
	<b>5</b>	4	6	4	7	5	3	5	4	7

Source: Authors' own elaboration on Inapp data

their educational qualifications not necessary for performing their current jobs. This must be related to a labour market scarcely oriented towards high professionalism, and also hints at a more general job dissatisfaction, not necessarily dependent on past educational choices.

Data shows a general tendency of reduction in overeducation for individuals with higher personal and household income: higher wages are typically associated with relevant professions with high human capital. The virtuous circle of high education-high wages-high job satisfaction seems to be still existing, even if the capacity of education to ensure

social mobility is being eroded in favour of the opposite relation: a good social position guarantees higher wages and gives the opportunity to attain high educational qualifications. The causality of the involved phenomena is anything but clear.

Is there a relationship between the channels of entry into employment and overeducation? Data shows that obtaining a job via a formal channel (e.g. via an open competitive selection or a recruiting company), or a professional channel gives a better match between qualification and job. In particular, a structured and professional placement determines a higher probability of getting a job more in line with

**Table 3b. Overeducation for several controls, %**

Job search channels		PES	Temporary agency	Research agency	Schools and Universities	Print ads	Professionals	Friends and relatives	Self nominations	Public competition	Own business
<b>1 Sheepskin effect</b>		39	48	25	19	36	32	48	36	14	35
<b>2 Overeducation</b>		24	29	13	9	18	13	27	19	7	16
<b>3 Overskill</b>		34	33	33	29	35	33	31	33	28	31
<b>4 Over ISCO</b>		31	42	17	15	21	18	27	23	9	20
<b>5 Over Mode</b>		18	29	21	22	18	16	16	20	12	15
<b>Factors of overeducation (1-5)</b>	<b>0</b>	33	26	37	48	37	38	31	34	55	38
	<b>1</b>	24	16	32	25	25	29	22	26	27	25
	<b>2</b>	13	16	16	13	16	16	16	17	9	17
	<b>3</b>	11	18	9	7	11	8	13	10	3	10
	<b>4</b>	8	14	2	3	4	4	8	6	2	4
	<b>5</b>	8	7	2	2	5	3	7	4	1	4
<b>Monthly household income (thousands)</b>											
		<1	1-2	2-3	3-5	>5	<b>Quintiles of gross yearly income</b>				
							I	II	III	IV	V
<b>1 Sheepskin effect</b>		47	38	31	25	26	43	40	32	25	24
<b>2 Overeducation</b>		29	20	16	12	11	25	21	18	12	11
<b>3 Overskill</b>		36	32	31	33	33	33	30	31	31	31
<b>4 Over ISCO</b>		26	23	19	17	13	21	23	23	20	16
<b>5 Over Mode</b>		14	16	18	19	17	16	16	17	18	16
<b>Factors of overeducation (1-5)</b>	<b>0</b>	31	37	42	45	45	34	38	41	44	46
	<b>1</b>	22	25	26	27	29	24	24	25	27	27
	<b>2</b>	19	15	14	13	14	18	15	13	13	15
	<b>3</b>	13	11	9	8	6	12	11	9	8	7
	<b>4</b>	8	7	5	4	3	6	6	6	5	3
	<b>5</b>	7	5	4	3	3	6	5	5	3	2

Source: Authors' own elaboration on Inapp data

own qualifications, rather than what happens with informal channels where recruiting is not finalized to maximize reciprocal satisfaction between demand and supply, but only rapid employment (Meliciani and Radicchia 2009; 2015). This finding has relevant policy implications because it highlights the higher probability of suffering some forms of mismatch when the channel of recruitment is a leasing company. Indeed, only 26% of those recruited in this way reported no factors of job mismatch, against the 55% of those participating in an open competitive selection. Also looking at the type of job contract, typical-atypical work and dependent-

self-employment, we note a significant penalisation in terms of being over-educated for atypical workers and temporary employees.

It seems apparent that negative judgements on job satisfaction are connected to a higher overeducation, reinforcing the interpretation that a bad matching fosters a growing work discomfort. The gap is a structural one: workers with higher work dissatisfaction report 10 extra percentage points of overeducation. The gap related to career prospects and economic treatment is less severe.

The mismatch may be transitory when it is the result of a process of integration of new

employees at the workplace. In this case, the mismatch is either reabsorbed over time through access to increasing roles, or the overeducation becomes systematic and chronic, and is a problem both at the individual level and at the level of the economic system, representing a cost. The economic interpretation is prevailing and a job that is at a lower level than that potentially adequate to the educational qualification acquired may respond to an explicit business strategy to save on the cost of younger employees now and recover in the future. This interpretation is not supported by the data: overeducation seems to be structural and a relevant share of individuals are mismatched even after many years of employment. Skills that are wasted often generate work discomfort, while a missing educational qualification is often superseded by the acquired professionalism.

#### 4. Methodology

In the econometric analysis, we enquire about the determinants and the wage penalty of overeducation. In this section, we summarize the methodology of both.

##### *The determinants of overeducation*

We study the determinants of overeducation by estimating probit models of the probability of experiencing overskilling and overeducation based on different individual and macroeconomic characteristics, with and without the Heckman sample selection correction procedure. The main equation is a probit of the determinants of being overeducated:

$$y_i^{o*} = X_i^o \gamma + \varepsilon_i^o \rightarrow y_i^o = I(y_i^{o*} > 0)$$

with  $\varepsilon_i^o \sim N(0, \sigma_\varepsilon^2)$ ,  $i=1, \dots, n$  [1]

where the apex "o" stands for overeducation or over-skilling,  $y$  is the outcome variable measuring the probability of being overeducated/skilled. It is not observed (latent variable), because it is just known if the individual is overeducated/skilled or not.  $X_i^o$  is a set of determinants of overeducation and the corresponding coefficients.  $I(\cdot)$  is an indicator function giving 1 if the condition holds and 0 otherwise.

Introducing the Heckman correction for the probability of being employed, we have:

$$z_i^{o*} = y_i^o \delta + u_i^o \quad \text{with} \quad u_i^o \sim N(0, \sigma_u^2) \quad [2]$$

where  $z_i^o = 1$  if  $z_i^{o*} > 0$   
 $z_i^o = 0$  if  $z_i^{o*} \leq 0$

and

$$\text{Prob}(z_i^o = 1) = \Phi(y_i^o \delta)$$

$$\text{Prob}(z_i^o = 0) = 1 - \Phi(y_i^o \delta)$$

where  $y_i^{o*}$  is observed only for those who are occupied, that is for  $z_i^o = 1$ .

The function estimate for the probability that an employed is overeducated is then corrected for the occupational selection term in the following manner:

$$E[y_i^o | X_i^o, z_i^o = 1] = X_i^o \gamma + E[\varepsilon_i^o | u_i^o > -y_i^o \delta] = X_i^o \gamma + \theta \lambda_i^o \quad [3]$$

where  $\lambda_i^o = \frac{\phi(y_i^o \delta)}{\Phi(y_i^o \delta)}$

is the inverse of the Mills ratio, with  $\phi(\cdot)$  and  $\Phi(\cdot)$ , respectively, probability density and cumulative standard normal distribution function.

In comparison to equation [1], equation [2] includes an additional set of covariates concerning the propensity to be employed. are unknown parameters. Indeed, through the Heckman correction, we test whether the above main equation is influenced by the probability to find employment. In other words, we ask whether those who are jobless would be overeducated or not if employed. To such an end, we add to [1] as a further regressor the inverse Mills ratio obtained through a selection equation estimated to predict the probability of employment. In this way, we test whether the jobless would be more or less likely to be overeducated. In the case of the positive correction coefficient, we have that taking into account the selection in employment increases the probability of being over-educated, which is a prevalence of the job search theory: those who are jobless have higher human capital characteristics and are in search of a well-matched job. They prefer to wait for unemployment rather than accept the first job offer they receive.

Conversely, in the case of a negative and significant correction coefficient, the selection in employment decreases the probability of being over-educated because those who are employed have higher human capital characteristics.

**Table 4. Control variables of model estimates**

Variables	Description
<i>Macro-region of residence</i>	<i>Reference category: South and Isles</i>
North West	North West=1, 0 otherwise
North East	North East= 1, 0 otherwise
Centre	Centre=1, 0 otherwise
<i>Age class:</i>	<i>Reference category: 18-29 years for high school graduated and 30-39 for tertiary educated</i>
30-39 years	30-39 Years=1, 0 otherwise
40-49 Years	40-49 years=1, 0 otherwise
50-64 Years	50-64 years=1, 0 otherwise
Woman	Gender: 1 woman; 0 men
Children	Presence of a child <14 years old: 1; no children 0
Woman with Children	Woman with at least one child =1, no=0
Numbers of income earners	Number of income earners in the household
Adult who lives in the parents' home	1 if the individual lives with their parents; 0 otherwise
Type_secondary	Type of secondary school education: 1="Liceo"; 0 other type of school (technical school)
<i>Failed</i>	<i>He repeated at least one class=1, no=0</i>
Diploma grade	Diploma mark
Delay in attaining graduation	1: yes; 0 otherwise
Graduation attempt	He attempted to attain tertiary education=1, no=0
Tertiary degree grade	Tertiary degree mark
Field of study in tertiary education:	Reference category: Law
Hard Sciences	Hard Sciences (Chemistry, Physics, Geology, Biology, Pharmacy, IT, Mathematics)=1, 0 otherwise
Medicine and Veterinary	Medicine and Veterinary=1, 0 otherwise
Engineering and Architecture	Engineering and Architecture =1, 0 otherwise
Soft Social Sciences	Soft Social Sciences (Sociology, Political Sciences, Communication, Sciences, Psychology) =1, 0 otherwise
Humanities	Humanities (Philosophy, Literature, Languages, Education) =1, 0 otherwise
Type of school attended	Public=1, 0 otherwise

Source: Authors' own elaboration on Inapp data

The variables included in equation [1] control for gender, the presence of children, age, type of high school diploma, the final grade at high school/university, the possible delay in attaining the highest level of education, the attempt to get a university degree for high secondary educated, knowledge of the English language, the attendance of a training course, the region of residence, the field of study and the population density of the area of residence.

The selection equation includes the same variables as the main equation, but also many instrumental variables, such as the number of household earners, if the individual still lives with

their parents and the unemployment rate. The detailed list of regressors is reported in Table 4.

#### ***The wage penalty of overeducation***

We adopt the usual Mincerian earnings equation augmented by a series of overeducation and/or overskilling dummies representing our above definitions as the empirical framework for estimating the wage penalty:

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

where  $\ln w_i$  is the natural logarithm of the net monthly



wage for individual  $i$ , the  $X_i$  are a set<sup>8</sup> of control variables assumed to affect earnings and the  $\beta_j$  are their coefficients. The  $O_i$  is a dummy equal to one when the individual  $i$  is mismatched and  $r^{OLS}$  is the estimated wage penalty associated with overeducation. The  $u_i$  is a disturbance term representing other forces which may not be explicitly measured, assumed independent of  $X_i$  and  $O_i$ . The latter dummy is here taken according to the specification adopted to mean overeducation, over-skilling or different interactions of the two or with other variables. The  $X$  variables include the same controls of previous equations [1] and [2], and more information about the type of job (job position, type of enterprise, economic activity, sector of employment). However, as this information can be endogenous, an alternative specification without these variables is also estimated. The detailed list of regressors is reported in Table 4.

Many observers have raised the concern that simple OLS estimates might tend to under/overestimate the wage penalty associated with the educational mismatch. There might be, in fact, unobserved heterogeneity between the overeducated and the rest of the sample of graduates (endogeneity bias) and, at the same time, between the employed and the non-employed (sample selection bias). As Kleibrink (2016) has recently shown, the first source of bias, the endogeneity bias, is generally addressed in three ways in the overeducation literature: 1) through longitudinal data, which however relies exclusively on the small group of those who change their educational matching and therefore is not very reliable in the case of overeducation (Korpi and Tåhlin 2009; Leuven and Oosterbeek 2011); 2) utilizing IV estimates, which, however, have also to face the lack of suitable instruments due to the strong link between the probability of being overeducated and the wage earned (Korpi and Tåhlin 2009); 3) by controlling for the quality of human capital in the regression, as Kleibrink himself does by using different indicators of skill coming from the German Socio-Economic Panel (GSOEP), and in particular the data from the International Adult Literacy Survey (IALS).

However, none of the previous strategies is easy to implement in our data. The panel dimension of the Inapp-PLUS data bank is too small to allow the first strategy. Moreover, the data set does not contain fully satisfactory instrumental variables and does not have variables able to measure skill and

talent as in the GSOEP/IALS data bank. Nonetheless, in what follows we try to test at least for sample selection bias, although some of the doubts existing for the instrumental variables used to correct for endogeneity hold also for sample selection bias.

Several economists have previously proposed the Heckman (1979) sample selection procedure (typically called Heckit for assonance to the Tobit model) as an empirical model to address the issue of omitted heterogeneity of the non-employed and assess its impact on the wage effect of overeducation. Neglecting the non-employed might generate a bias on returns to education and also on the wage effect of the educational mismatch whose direction is in principle ambiguous (see, among others, Sloane *et al.* 1999; Dolton and Vignoles 2000; Cuttillo and Di Pietro 2006).

We follow Caroleo and Pastore (2018) in extending the approach of Nicaise (2001) who suggested interpreting the sign of the bias in the Heckit estimated coefficient, as compared to the OLS estimates, as an indication in support for one of the two alternative theories of unemployment, namely what he calls the “crowding hypothesis” and the “reservation wage” hypothesis based on the job search model. Caroleo and Pastore (2018, Figure 1) apply this line of reasoning to the case of overeducation. They show that the Heckit can be used as a screening device to select among alternative theoretical interpretations of overeducation. In the job competition, the job assignment and the human capital model, sample selection bias might arise since the mismatch appears at first in the form of a higher probability of non-employment and only at a later stage it takes the form of a wage penalty for overeducation. After controlling for the selection bias arising from considering the non-employed, hence, the wage penalty associated with overeducation should be higher. Conversely, according to the job search theoretical model, non-employment is a voluntary choice, since the most skilled graduates prefer to remain non-employed while waiting for a better job offer to come. If employed, they would be less likely to experience overeducation. In this case, after controlling for the selection bias arising from considering non-employment, the wage penalty of overeducation should be lower.

The Heckit specification of the earnings equation suggests a method to correct the usual OLS estimates for the lower/higher employment opportunities of

8 We exclude job characteristics, partly endogenous.

the most skilled and motivated among those whose personal attributes would lead to overeducation if they were employed. In analytical terms, equation [1] should be specified differently:

$$\ln w_i = r^{\text{Heckit}} \theta_i + \sum_{j=1}^n \beta_j X_{i,j} + \rho \lambda \left( \sum_{l=1}^m \theta_l Z_{i,l} \right) + u_i \quad [2]$$

where  $r$  is now denoted with the superscript Heckit, to distinguish it from the corresponding OLS estimate;  $\rho$  is the correlation between the error terms of the main and of the participation equation and  $\lambda$  is the inverse Mills ratio evaluated at the mean of the covariates ( $Z$ ), which includes, 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.

Two possibilities are in order:

$$\begin{aligned} H_0: r^{\text{Heckit}} &> r^{\text{OLS}} \\ H_1: r^{\text{Heckit}} &\leq r^{\text{OLS}} \end{aligned} \quad [3]$$

As shown in of Caroleo and Pastore (2018, Figure 1, panel (a)), according to  $H_0$ , OLS is underestimating the wage penalty associated with overeducation/overskilling. Only the most skilled overeducated are selected for employment.  $H_0$  is consistent with the job competition and also the job assignment model<sup>9</sup>, 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.  $H_0$  is also consistent with the human capital model, in as much as selection into employment refers to the graduates with the highest quality of human capital. The least skilled would be such because of their lack of work-related competencies: their lower human capital tends to relegate them to non-employment.

According to  $H_1$ , OLS is overestimating the wage penalty associated with the educational mismatch. Only the least skilled and overeducated are selected into employment.  $H_1$  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 non-

employment pool for the best job offer to arrive.

The variables used in the selection equation are finalized to identify the different probabilities of individual employability. We include two sets of variables. Set one includes the same determinants of the main equation and set two includes the instrumental variables. We use the unemployment rate, the presence of a university and the use of recruitment agencies in the area of residence. Finally, in some of the estimates, we also use information about the public or private nature of the school attended by the respondent. We expect that these instrumental variables follow the exclusion restriction, namely that they affect the probability to find a job, but not the type of wage earned. We test this hypothesis in the empirical analysis.

## 5. Findings

We see the results found both on the side of the measure of the phenomenon and on that of the salary penalty procured.

### *The determinants*

Tables 5 and 6 report the Heckprobit estimates of the determinants of the educational mismatch respectively for upper secondary school graduates (Table 5) and tertiary graduates (Table 6). As it should be now clear, the Heckprobit is a two-equation model where the main equation is a probit of being overeducated and the selection equation is a probit of being employed. Each table shows different empirical models, one for each definition of overeducation/overskilling. Model (1) refers to the probability of being overskilled conditional on being employed (Model 2). Model (3) estimates the probability of being overeducated in comparison to the declared level of education, always conditional on being employed. Models (5) and (6) measure the determinants of overeducation related to the level of education needed for the specific job. Models (7) and (8) refer to the probability of being overeducated according to the ISCO approach.

In Table 5, the variables used as instruments in the selection equation are the unemployment rate of the province of residence (nuts3 level detail), the number of household's income members which act by increasing the probability of being

<sup>9</sup> We thank Peter Sloane for suggesting this extension to us.

**Table 5. Heckprobit estimates of the determinants of the educational mismatch among high secondary school graduates**

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Skill mismatch Main equation	Prob employed Selection equation	Educational mismatch 1 (declared) Main equation	Prob employed Selection equation	Educational mismatch 2 Main equation	Prob employed Selection equation	ISCO mismatch Main equation	Prob employed Selection equation
Woman	-0.0177 (0.0694)	-0.387*** (0.0366)	-0.159*** (0.0587)	-0.387*** (0.0365)	-0.191*** (0.0509)	-0.388*** (0.0365)	-0.806*** (0.0656)	-0.378*** (0.0367)
Children	0.125* (0.0747)	0.00620 (0.0441)	-0.117* (0.0698)	0.00476 (0.0441)	-0.130** (0.0576)	0.00156 (0.0442)	-0.0795 (0.0644)	-0.00181 (0.0444)
Woman with children	-0.202** (0.0917)	-0.411*** (0.0490)	0.00636 (0.0812)	-0.411*** (0.0490)	-0.0746 (0.0679)	-0.411*** (0.0490)	0.0599 (0.0871)	-0.406*** (0.0493)
Ttype_of_Secondary School (ref. "liceo")	0.0136 (0.0663)	-0.115*** (0.0331)	0.0809 (0.0575)	-0.115*** (0.0331)	0.401*** (0.0502)	-0.112*** (0.0332)	-0.167** (0.0678)	-0.114*** (0.0331)
Failed one year of school	0.0803 (0.0580)	-0.109*** (0.0294)	-0.0142 (0.0520)	-0.109*** (0.0294)	0.0771* (0.0444)	-0.109*** (0.0294)	0.0459 (0.0520)	-0.108*** (0.0295)
Diploma grade	-0.00377** (0.00187)	0.00335*** (0.000957)	-0.00950*** (0.00182)	0.00332*** (0.000958)	-0.00690*** (0.00143)	0.00336*** (0.000958)	-0.00767*** (0.00182)	0.00339*** (0.000961)
Graduation attempt	0.140*** (0.0527)	0.131*** (0.0273)	-0.116** (0.0483)	0.131*** (0.0273)	-0.0364 (0.0402)	0.130*** (0.0273)	-0.147*** (0.0506)	0.132*** (0.0274)
He knows English and PC science	0.231 (0.160)	0.233*** (0.0662)	-0.455*** (0.115)	0.236*** (0.0663)	-0.468*** (0.110)	0.236*** (0.0662)	-0.639*** (0.111)	0.227*** (0.0663)
He did a training course	0.113** (0.0549)	0.618*** (0.0267)	-0.390*** (0.0502)	0.618*** (0.0267)	-0.298*** (0.0418)	0.619*** (0.0267)	-0.260*** (0.0514)	0.622*** (0.0268)
Area of residence (ref. South)								
North West	-0.137** (0.0676)	0.481*** (0.0394)	-0.0162 (0.0646)	0.479*** (0.0394)	0.00209 (0.0521)	0.481*** (0.0393)	0.175*** (0.0641)	0.489*** (0.0396)
North East	-0.164** (0.0731)	0.492*** (0.0389)	0.0538 (0.0653)	0.491*** (0.0388)	0.0427 (0.0539)	0.494*** (0.0388)	0.265*** (0.0648)	0.502*** (0.0390)
Center	-0.0490 (0.0700)	0.356*** (0.0383)	0.0588 (0.0623)	0.354*** (0.0384)	0.105** (0.0516)	0.354*** (0.0382)	0.198*** (0.0637)	0.360*** (0.0385)
Age class (ref. 18-29 years)								
30-39 years	0.116* (0.0660)	0.558*** (0.0360)	0.0746 (0.0586)	0.556*** (0.0361)	0.186*** (0.0506)	0.554*** (0.0360)	0.170*** (0.0594)	0.563*** (0.0361)
40-49 years	0.0433 (0.0759)	0.686*** (0.0404)	-0.204*** (0.0694)	0.684*** (0.0406)	0.000883 (0.0574)	0.682*** (0.0405)	-0.0680 (0.0699)	0.695*** (0.0406)
50-64 years	0.0349 (0.0688)	0.168*** (0.0393)	-0.279*** (0.0612)	0.166*** (0.0394)	-0.251*** (0.0516)	0.166*** (0.0393)	-0.289*** (0.0632)	0.179*** (0.0394)
Unemployment rate		-0.00176 (0.00274)		-0.00182 (0.00274)		-0.00174 (0.00273)		-0.00135 (0.00276)
N. of earners income		0.791*** (0.0224)		0.791*** (0.0224)		0.790*** (0.0225)		0.794*** (0.0225)
Adult who lives in the parents' home		-0.791*** (0.0404)		-0.797*** (0.0412)		-0.803*** (0.0409)		-0.788*** (0.0405)
Constant	-1.352*** (0.243)	-1.492*** (0.119)	0.497** (0.206)	-1.487*** (0.120)	0.798*** (0.178)	-1.486*** (0.119)	0.640*** (0.208)	-1.520*** (0.120)
At rho		-8.62e-05 (0.106)		0.0824 (0.0879)		0.154** (0.0783)		0.200** (0.102)
Observations	16,281	16,821	16,281	16,281	16,281	16,281	16,224	16,224

Note: Robust standard errors in parentheses \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1.

Source: Authors' own elaboration on Inapp data

employed and the condition of living with their parents which, instead, acts in opposite direction, as expected because when a young person reaches a certain degree of financial autonomy, (s)he will be more likely to decide to live alone. In omitted estimates, we test that these variables do respect the exclusion restriction, namely that they do not influence the probability of being overeducated. Now we test whether they influence the probability of finding a job. In all the selection equations, the unemployment rate is not statistically significant, while the number of earners affects positively the probability of finding a job, with a coefficient of 0.79 in all the models, confirming the presence of strong informal networks in the Italian labour market, while the presence of adults in the house reduces the employment probability, most likely of women, who are more likely to take care of the dependent people within the household. The low statistical significance of the unemployment rate probably depends on the presence in the estimates of dummies for the macro-regions of residence which already absorb the effect of the local labour market conditions on employment chances.

It is important to note that the coefficient of the Mills ratio in the main equation, namely the *athrho*, is statistically significant only in models (6) and (8), denoting that the simple OLS estimates are biased in these cases. Coefficients are 0.15 and 0.20 respectively in models (6) and (8) and their positive sign is following the job competition, job assignment and human capital theories. This means that the probability of being over-educated increases after that we have considered the selection effect in employment: individuals may choose to accept jobs for which they are over-qualified to accumulate work experience and enhance future career prospects. This evidence is also in line with the scarce job opportunities typical of labour markets with high levels of (youth and adult) unemployment, such as the Italian one.

For secondary high school graduates, women tend to have a lower probability of being overeducated, but not overskilled. The skill-mismatch increases in presence of a child. However, this latter outcome seems to be effective only for men, because being a woman and having a child significantly reduces the probability of being overskilled. The penalty for being a woman with children on the probability of being employed is very high: indeed, in all the models the coefficients are lower than -0.40.

Having attained the diploma with a higher grade significantly increases the probability of being employed and decreases the probability of being mismatched in skills or in education while having repeated one class reduces the probability of being employed and, in one case, the probability of being overeducated, but does not seem to affect the probability of overskilling.

When we consider the effects of activities finalized to increase human skills, such as the acquisition of skills in personal computer use and better English knowledge or the attendance of a training course, we can see that they all increase the probability of being employed (the coefficients are, respectively, around the 0.23 and 0.62), reduce the probability of being over-educated, increase the probability of being over-skilled, as they increase the personal human capital without increasing the level of education.

Regarding the macro-region of residence, living in the Centre or in the North of Italy acts positively on the probability of being employed and decreases that of over-skilling. In comparison to living in the South of Italy, the probability of being employed increases of 0.36 for the Centre, 0.48 for the Northwest and 0.49 for the Northeast while the probability of being overskilled reduces by 0.05 for the Centre, 0.14 for the North West and 0.16 for the North East. Conversely, regarding overeducation, the residence in the Centre or the North increases the probability to experience overeducation, but only according to the ISCO definition, with coefficients between 0.17 and 0.29.

Finally, according to the age class, being older increases the probability of being employed and decreases that of being over-educated but has no statistically significant relationship with the probability to experience overskilling.

As regards the tertiary graduates (Table 6), the determinants of educational and skill mismatch are very similar to those relative to high secondary school holders. About the field of study, in comparison to a degree in law, having attained a university degree in soft social sciences or in humanities determines a significant reduction in the probability of being employed while a degree in Medicine and Veterinary or in Engineering and Architecture significantly reduces the probability of being over-educated (according to the ISCO classification, the

**Table 6. Heckprobit estimates of the determinants of the educational mismatch among university graduates**

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Skill mismatch Main equation	Prob employed Selection equation	Educational mismatch 1 (declared) Main equation	Prob employed Selection equation	Educational mismatch 2 Main equation	Prob employed Selection equation	ISCO mismatch Main equation	Prob employed Selection equation
Woman	0.0252 -0.0711	-0.255*** -0.0404	-0.0129 -0.0669	-0.254*** -0.0404	-0.108* -0.0644	-0.255*** -0.0404	0.0831 -0.0639	-0.250*** -0.0404
Children	0.0802 -0.0867	-0.0425 -0.0564	-0.123 -0.085	-0.0404 -0.0565	-0.05 -0.0808	-0.0415 -0.0565	-0.0179 -0.0812	-0.0428 -0.0564
Woman with children	-0.224** -0.0978	-0.282*** -0.0612	-0.0585 -0.0954	-0.283*** -0.0612	-0.0761 -0.0916	-0.283*** -0.0613	-0.0558 -0.0907	-0.281*** -0.0612
Type secondary (ref. "Liceo")	0.0192 -0.0513	-0.00111 -0.0317	-0.148*** -0.0488	-0.00164 -0.0316	-0.136*** -0.0477	-0.00198 -0.0317	-0.152*** -0.047	0.00043 -0.0316
Failed one year of school	-0.126 -0.089	-0.107** -0.0531	0.0775 -0.0827	-0.105** -0.0531	0.0998 -0.0812	-0.107** -0.0532	-0.0217 -0.0829	-0.102* -0.0531
Diploma grade	-0.00102 -0.00149	-0.00115 -0.00089	-0.000114 -0.00142	-0.00113 -0.00089	-4.28E-05 -0.0014	-0.00115 -0.00089	0.000411 -0.00133	-0.00116 -0.00089
University grade	-0.00546 -0.00379	0.00741*** -0.00213	-0.0187*** -0.00338	0.00758*** -0.00212	-0.0146*** -0.00329	0.00746*** -0.00212	-0.00815** -0.00328	0.00750*** -0.00212
Dealy in graduation	0.0152 -0.0527	0.0373 -0.0308	-0.106** -0.0494	0.0355 -0.0307	-0.168*** -0.0479	0.0363 -0.0307	-0.146*** -0.0458	0.0375 -0.0307
<i>Fields of study (ref. Law)</i>								
Hard Sciences	-0.00898 -0.0818	0.0257 -0.0483	-0.374*** -0.0791	0.0254 -0.0483	-0.311*** -0.0768	0.0258 -0.0484	-0.620*** -0.0724	0.0252 -0.0483
Medicine and Veterinary	-0.0343 -0.0856	0.0737 -0.0556	-0.529*** -0.0933	0.0739 -0.0555	-0.590*** -0.091	0.0741 -0.0556	-0.932*** -0.0863	0.0753 -0.0556
Engineering and Architecture	0.00711 -0.0811	-0.00184 -0.0482	-0.393*** -0.0804	-0.0016 -0.0482	-0.272*** -0.0765	-0.00193 -0.0485	-0.572*** -0.0737	0.000752 -0.0481
Soft Social Sciences	0.179** -0.0867	-0.131** -0.0549	0.392*** -0.0788	-0.131** -0.0549	0.481*** -0.0772	-0.130** -0.0549	0.0346 -0.0752	-0.132** -0.0548
Humanities	0.0103 -0.0743	-0.141*** -0.0434	0.0792 -0.066	-0.140*** -0.0434	0.143** -0.065	-0.140*** -0.0434	-0.400*** -0.0638	-0.141*** -0.0434
He knows English and PC science	0.00901 -0.339	0.437** -0.173	0.281 -0.413	0.434** -0.173	-0.403 -0.289	0.436** -0.173	0.318 -0.287	0.430** -0.175
He did a training course	0.160*** -0.0558	0.563*** -0.0288	-0.355*** -0.0526	0.560*** -0.0288	-0.257*** -0.0518	0.562*** -0.0288	-0.361*** -0.0493	0.561*** -0.0288
At least a high school graduate parent	-0.0291 -0.0522	-0.0813** -0.0317	-0.0559 -0.0492	-0.0823*** -0.0317	0.0820* -0.048	-0.0813** -0.0317	0.0297 -0.0472	-0.0809** -0.0317
<i>Area of residence (ref. South)</i>								
North West	-0.190*** -0.0713	0.446*** -0.04	0.0436 -0.0687	0.446*** -0.04	0.0828 -0.0655	0.446*** -0.04	0.0786 -0.0642	0.447*** -0.04
North East	-0.145** -0.0721	0.377*** -0.041	0.135* -0.0699	0.376*** -0.041	0.257*** -0.0661	0.376*** -0.041	0.109 -0.067	0.376*** -0.0409
Center	-0.0407 -0.065	0.337*** -0.0393	0.00288 -0.0656	0.339*** -0.0395	0.0932 -0.0631	0.336*** -0.0393	0.162*** -0.0623	0.338*** -0.0393
<i>Age class (ref. 30-39 year)</i>								
18-24 years	-0.166 -0.137	-1.491*** -0.059	0.0841 -0.129	-1.490*** -0.059	-0.00687 -0.122	-1.490*** -0.059	0.271** -0.119	-1.490*** -0.059
25-29 years	-0.201*** -0.0754	-0.861*** -0.0405	-0.0837 -0.0716	-0.861*** -0.0405	-0.0479 -0.0677	-0.860*** -0.0405	0.251*** -0.0648	-0.860*** -0.0404
40-49 years	-0.156** -0.0752	0.216*** -0.0497	-0.225*** -0.0699	0.216*** -0.0496	-0.221*** -0.0669	0.215*** -0.0496	-0.178*** -0.0649	0.217*** -0.0496
50-64 years	-0.190*** -0.0688	-0.251*** -0.0452	-0.400*** -0.0681	-0.251*** -0.0452	-0.385*** -0.0658	-0.252*** -0.0452	-0.340*** -0.0648	-0.248*** -0.0452

continued on next page



Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Skill mismatch Main equation	Prob employed Selection equation	Educational mismatch 1 (declared) Main equation	Prob employed Selection equation	Educational mismatch 2 Main equation	Prob employed Selection equation	ISCO mismatch Main equation	Prob employed Selection equation
<i>Degree of urbanisation (ref. &gt;250,000 inhabitants)</i>								
< 5,000 inhabitants	-0.0543 -0.0823	0.0269 -0.0492	-0.0583 -0.0786	0.0244 -0.0491	-0.0828 -0.0758	0.0255 -0.049	-0.0581 -0.0737	0.0195 -0.0492
5,001 - 20,000 inhabitants	-0.037 -0.0734	0.0615 -0.0434	-0.0305 -0.0689	0.0578 -0.0435	-0.0923 -0.0669	0.0608 -0.0434	0.0174 -0.0656	0.0575 -0.0434
20,001 - 50,000 inhabitants	-0.0227 -0.0772	0.0595 -0.0471	0.054 -0.0732	0.058 -0.0472	-0.0834 -0.0714	0.0589 -0.0471	-0.0916 -0.0706	0.0584 -0.0471
50,001 - 250,000 inhabitants	-0.187** -0.0728	0.0877** -0.0442	-0.131* -0.0706	0.0856* -0.0441	-0.174** -0.0679	0.0863* -0.0441	-0.042 -0.0659	0.0867** -0.0441
Number of income earners		0.644*** -0.0262		0.645*** -0.0262		0.644*** -0.0263		0.646*** -0.0261
Adult who lives in the parents' home		-0.951*** -0.0466		-0.949*** -0.0469		-0.952*** -0.0468		-0.945*** -0.0469
Constant	-0.359 -0.533	-1.260*** -0.281	1.484*** -0.555	-1.276*** -0.281	1.651*** -0.462	-1.262*** -0.281	0.497 -0.453	-1.272*** -0.282
athrho		0.0875 -0.128		-0.118 -0.114		0.0131 -0.11		-0.229** -0.106
<b>Observations</b>	<b>11276</b>	<b>11276</b>	<b>11276</b>	<b>11276</b>	<b>11276</b>	<b>11276</b>	<b>11268</b>	<b>11268</b>

Note: Robust standard errors in parentheses \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Source: Authors' own elaboration on Inapp data

coefficient for overeducation is -0.93), but does not affect the probability of being over-skilled. Having at least a parent with a high school degree shows a coefficient of -0.08 in all the models, denoting a reduction in the probability of being employed, probably because in this case young people are more oriented to continue their studies and do not accept any type of job. Finally, considering the degree of urbanisation of the place of residence, in comparison to the city, only living in a medium-sized town (with 50,000-250,000 inhabitants) significantly reduces the probability of being mismatched and increases that of being employed, probably for the activation of the friends and relatives network typical of small towns. In turn, this confirms that in Italy informal networks still represent the most used channel to find a job (Meliciani and Radicchia 2009).

The athrho coefficient is statistically significant only in model (8) (overeducation based on the ISCO classification). The negative sign shows that accounting for the instrumental variables, which, in this case, are the information on the number of earners in the household and the coexistence with parents, the estimated probability of being

overeducated reduces. Also, for tertiary educated, living with their parents significantly decreases the probability of being employed while the number of income earners acts increasing it.

In conclusion, the family background seems to play a relevant role in the probability of being over-educated: higher parents' level of education, a higher number of family members and living with the parents' home result indeed all factors reducing the probability of being over-educated. However, despite what happens for upper secondary school graduates, for tertiary educated these personal instrumental factors to act reducing the probability of being over-educated, at least according to the ISCO definition. Probably, their higher human capital characteristics tend to reduce their attitude to accept any type of job. They are therefore more prone to wait for the most appropriate job according to their educational background.

### The wage penalty

Finally, Table 7 reports the results of estimates of the wage penalty due to overeducation, overskilling and the double penalisation of being both overeducated and overskilled. Also, in this case, different empirical

**Table 7. Earning equations with and without occupational covariates. Wage mismatch determinants. Different models specification for Upper secondary and University graduates**

Variables	Upper secondary graduates				University graduates			
	OLS	OLS	Heckit		OLS	OLS	Heckit	
	Earning equation With occupational covariates	Earning equation Without occupational covariates	Earning equation Main equation	Employment equation Heckman	Earning equation With occupational covariates	Earning equation Without occupational covariates	Earning equation Main equation	Employment equation Heckman
Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Educational and skill mismatch (cat. Matched)	-0.137*** (0.0461)	-0.170*** (0.0458)	-0.137*** (0.0355)		-0.0924** (0.0454)	-0.125*** (0.0439)	-0.132*** (0.0511)	
Only educational mismatch (declared)	-0.0844*** (0.0243)	-0.110*** (0.0236)	-0.106*** (0.0357)		-0.0505* (0.0278)	-0.0881*** (0.0266)	-0.0605** (0.0277)	
Only skill mismatch	0.0458* (0.0264)	0.0438* (0.0264)	0.0306 (0.0253)		-0.0119 (0.0305)	-0.00983 (0.0305)	0.00192 (0.0305)	
Woman	-0.120*** (0.0224)	-0.110*** (0.0218)	-0.112*** (0.0263)	-0.471*** (0.0427)	-0.103*** (0.0276)	-0.111*** (0.0276)	-0.118*** (0.0304)	-0.249*** (0.0497)
Children	0.0854*** (0.0203)	0.0951*** (0.0202)	0.105*** (0.0242)	0.283*** (0.0450)	0.0812*** (0.0287)	0.0879*** (0.0288)	0.0940** (0.0401)	0.499*** (0.0618)
Woman with children	-0.0468* (0.0276)	-0.0483* (0.0273)	-0.0690** (0.0319)	-0.423*** (0.0556)	-0.0909*** (0.0347)	-0.0891** (0.0348)	-0.0908** (0.0436)	-0.293*** (0.0707)
Type_secondary (ref. "liceo")	0.0126 (0.0201)	0.00747 (0.0200)	0.0195 (0.0196)	-0.243*** (0.0399)				
Failed one or more years of school	0.0266 (0.0166)	0.0251 (0.0166)	0.0183 (0.0193)	-0.182*** (0.0340)				
Diploma grade	0.00178*** (0.000519)	0.00193*** (0.000517)	0.00179*** (0.000625)	0.00472*** (0.00108)				
Graduation attempt	0.0473*** (0.0149)	0.0560*** (0.0147)	0.0454*** (0.0160)	0.129*** (0.0310)				
He knows English and PC science	0.0165 (0.0449)	0.0473 (0.0448)	0.0852* (0.0476)	0.367*** (0.0801)	0.0143 (0.110)	0.00503 (0.111)	0.0299 (0.0568)	0.385* (0.211)
He did a training course	0.0706*** (0.0137)	0.0762*** (0.0135)	0.0743*** (0.0159)	0.745*** (0.0297)	0.0900*** (0.0174)	0.0867*** (0.0174)	0.102*** (0.0278)	0.659*** (0.0344)
<i>Area of residence (ref. South)</i>								
North West	0.0717*** (0.0239)	0.0642*** (0.0237)	0.0953*** (0.0224)	0.659*** (0.0493)	0.145*** (0.0317)	0.141*** (0.0318)	0.138*** (0.0267)	0.564*** (0.0680)
North East	0.116*** (0.0212)	0.110*** (0.0211)	0.131*** (0.0233)	0.733*** (0.0434)	0.161*** (0.0286)	0.154*** (0.0287)	0.149*** (0.0308)	0.502*** (0.0589)
Center	0.0559** (0.0227)	0.0526** (0.0226)	0.0629*** (0.0213)	0.498*** (0.0473)	0.0337 (0.0298)	0.0269 (0.0299)	0.0548* (0.0328)	0.394*** (0.0605)
<i>Age class (ref. 18-29 years)</i>								
30-39 years	0.0506** (0.0218)	0.0549** (0.0217)	0.0480** (0.0244)	0.763*** (0.0417)	0.0236 (0.0263)	0.0290 (0.0264)	0.0519 (0.0355)	1.276*** (0.0455)
40-49 years	0.133*** (0.0225)	0.139*** (0.0224)	0.143*** (0.0239)	0.959*** (0.0449)	0.151*** (0.0304)	0.173*** (0.0303)	0.230*** (0.0461)	1.717*** (0.0605)
50-64 years	0.223*** (0.0207)	0.232*** (0.0202)	0.238*** (0.0223)	0.681*** (0.0427)	0.297*** (0.0275)	0.333*** (0.0266)	0.375*** (0.0409)	1.497*** (0.0528)
University grade (ref. 110/110 with honneurs)					0.0493*** (0.0178)	0.0583*** (0.0178)	0.0741*** (0.0244)	0.211*** (0.0375)
Delay in attaining graduation					-0.0650**	-0.0709**	-0.0508*	-0.194***
<i>Fields of study (ref. Humanistic area)</i>								
Hard Sciences					0.0720*** (0.0253)	0.0728*** (0.0252)	0.0842*** (0.0264)	0.220*** (0.0520)
Medicine and Veterinary					0.284*** (0.0298)	0.235*** (0.0292)	0.248*** (0.0336)	0.354*** (0.0647)

continued on next page

Variables	Upper secondary graduates				University graduates			
	OLS	OLS	Heckit		OLS	OLS	Heckit	
	Earning equation With occupational covariates	Earning equation Without occupational covariates	Earning equation Main equation	Employment equation Heckman	Earning equation With occupational covariates	Earning equation Without occupational covariates	Earning equation Main equation	Employment equation Heckman
Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Engineering area					0.117*** (0.0345)	0.129*** (0.0331)	0.200*** (0.0482)	0.237*** (0.0676)
Architecture area					0.0235 (0.0550)	0.0337 (0.0553)	0.0310 (0.0453)	-0.236** (0.101)
Economic- statistics area					0.0838*** (0.0304)	0.0455 (0.0289)	0.0763* (0.0397)	0.226*** (0.0593)
Political-social area					0.0530 (0.0362)	0.0269 (0.0356)	0.0125 (0.0478)	-0.000861 (0.0707)
Legal area					0.142*** (0.0385)	0.125*** (0.0379)	0.117** (0.0465)	-0.135* (0.0753)
<i>Professional qualification (ref. High)</i>								
Medium	-0.0967*** (0.0185)				-0.140*** (0.0207)			
Low	-0.156*** (0.0262)				-0.0859 (0.0941)			
Type of enterprise (ref. private)	-0.00776 (0.0176)				-0.0895*** (0.0244)			
<i>Economic activity sector (ref. Production)</i>								
Production services	0.00598 (0.0218)				-0.0806** (0.0366)			
Distribution services	-0.000416 (0.0217)				-0.00862 (0.0408)			
Personal services	-0.0319 (0.0270)				-0.00455 (0.0497)			
Social services	-0.0414* (0.0232)				-0.150*** (0.0367)			
Unemployment rate	-0.00123 (0.00156)	-0.00170 (0.00156)		-0.00468 (0.00316)	0.00179 (0.00204)	0.00191 (0.00204)		-0.00305 (0.00418)
Presence of University	-0.00105 (0.00132)	-0.00123 (0.00131)		-0.000238 (0.00272)	0.00299* (0.00164)	0.00293* (0.00165)		0.000256 (0.00340)
Presence of PES and Employment services	9.90e-05 (6.07e-05)	0.000107* (6.05e-05)		3.12e-05 (0.000131)	-7.66e-05 (7.86e-05)	-9.36e-05 (7.88e-05)		-9.90e-05 (0.000169)
Type of school attended (ref. public)	-0.00359 (0.0320)	-0.00897 (0.0319)		0.0234 (0.0636)	0.0652* (0.0388)	0.0670* (0.0390)		0.164** (0.0807)
Constant	9.800*** (0.0801)	9.649*** (0.0752)	9.580*** (0.0720)	-1.443*** (0.143)	9.988*** (0.131)	9.807*** (0.127)	9.776*** (0.0894)	-2.099*** (0.245)
athrho			0.0186 (0.0239)				0.103* (0.0579)	
Insigma			-0.823*** (0.0416)				-0.698*** (0.0505)	
<b>Observations</b>	<b>4,693</b>	<b>4,756</b>	<b>11,942</b>	<b>11,942</b>	<b>3,484</b>	<b>3,507</b>	<b>7,704</b>	<b>7,704</b>
R-squared	0.126	0.119			0.192	0.174		

Note: Robust standard errors in parentheses \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Source: Authors' own elaboration on Inapp data

models are presented. The first model is an augmented specification including, in addition to the variables already included in Tables 5 and 6, information on the type of job, that is the professional qualification, the

type of enterprise (private/public), the economic activity, the sector and a set of variables controlling for the economic and social context where the individual has grown (regional unemployment rate, presence of

University in the area of residence, of Public Employment Services and private Employment agencies, and the type, public or private, of the school attended). Model (2) differs from model (1) for the exclusion of the covariates concerning the type of job while model (3) omits also the information linked to the economic and social context. Finally, model (4) estimates the expected wage conditional to the probability of being employed through the Heckman correction. Models 5-8 are the correspondent ones for tertiary educated. The coefficients remain substantially stable within the different models, highlighting overall that the determinants of wages are very similar between upper secondary graduates and tertiary educated.

Overeducation, even when combined with overskilling, significantly reduces wages. The penalisation due only to the condition of overeducation is more severe for upper secondary graduates, with coefficients ranging from -0.08 in model (1) to -0.11 in model (2) while for university graduates, they range from -0.05 in model (1) to 0.09 in model (2). The double penalisation due to the contemporary condition of overeducation and overskilling still increases the wage penalisation, with coefficients which reach -0.17 in model (2) for upper secondary graduates and -0.125 for tertiary graduates. Conversely, being only overskilled tends to increase the wage, but only for upper secondary school graduates. The regression coefficient is significant only in models (1) and (2,) with values of, respectively, 0.0458 and 0.0438. This is a good result showing that, despite the level of education, meritocracy tends to prevail at least for upper secondary education.

Being a woman, with or without children, significantly reduces the wage and the probability of being employed. Passing from upper secondary education to a university degree significantly reduces the gender gap in the probability of being employed (from -0.471 to -0.249) while the reduction in the wage gap is small and detected only in correspondence of model (1), where the coefficient passes from -0.12 to -0.10 while for the other models, they are stable at 0.11. According to the place of residence, in comparison to the South of Italy left as the reference category, living in the Centre or the North of Italy produces a significant increase in wages and in the probability of being employed. The earning premium increases for the North in comparison to the Centre, for the Northeast the effect is stronger for tertiary educated, in comparison to upper secondary graduates. Indeed, for upper secondary graduates, the re-

gression coefficients move from values around 0.05 for the Centre to values around 0.12 for the Northeast; for university graduates the regression coefficients move from values around 0.03 for the Centre to 0.15 for the Northeast. Wages increase also with age, as expected, according to the human capital theory, as the individual tends to accumulate more years of work experience.

About the education attained, for those with no tertiary education a high diploma mark significantly increases their wage; the same result is also connected to the attempt to attain a tertiary degree. The correspondent regression coefficient is indeed significant showing a wage premium of 0.05 while the coefficient for the probability of being employed is 0.129. Evidently, the experience lived at the University, even when the path is not concluded, significantly increases the individuals' employability in terms of competencies acquired, also if not certified. Conversely, for tertiary educated, any field of study different from the humanistic one produces a significant increase in wages earned since the higher probability of being employed, with the only exception of architecture and social and political sciences. The maximum wage increase is related to the medical area, with a coefficient of 0.28 according to model (1), 0.235 for model (2) and 0.248 for model (3). It is followed by law and engineering sciences. Other characteristics connected with higher wages are a high diploma mark and having attended graduation without a delay.

As for the characteristics of the job, higher professional qualifications and working in the public sector significantly increase wages. The wage penalty associated with a job in the private sector is significant only for tertiary educated, with a coefficient of -0.0895 while, among the economic activity sectors, the only one associated with a significant reduction in wages when the production sector has been used as reference category is social services (-0.04 for upper secondary educated and -0.15 for tertiary educated).

Finally, focusing on the economic and social context, while the presence of a university in the place of residence and having attended a public school significantly increases the wages only for tertiary educated (with regression coefficients respectively of 0.0029 and 0.067), the presence of PES (Public Employment Service) and private employment agencies seems to act in the direction of increasing the wages only for the upper secondary educated.

The coefficient for the Heckman correction highlights the presence of a significant sample selection

only for tertiary educated. The positive sign of the coefficient denotes, that correcting for the selection effect in employment, the expected income increases. Higher wages are associated with those who are employed. Therefore, in conclusion, concerning tertiary educated, according to the results of Table 6, there is a negative selection effect for the probability of being over-educated, which means that those who tend to be employed soon are more likely to be over-educated. However, there is also a positive selection effect on wages, denoting that once a university graduate finds a job it is more likely that it will be a well-paid and good-matched job. In other terms, in comparison to the lower educated, for the tertiary educated the time spent before finding a job seems more finalized to the search for a more adequate job.

Comparing the OLS coefficients with Heckit as described in the methodology, we conclude that for upper secondary graduates the wage penalty due to overeducation is stronger when we do not account for the Heckman correction. This is consistent with the job competition, job assignment and human capital model. For university graduates, the OLS and the Heckit coefficients are very similar, but the Heckit ones are higher, which is more in line with the job search model. This may be due to the lower degree of unemployment among adult workers and the greater substitutability of adult workers between insiders and outsiders. However, the non-statistical significance of the  $\alpha$  coefficient suggests caution in assuming such a conclusion.

### Conclusions and policy suggestions

Starting in 2014, the Inapp-PLUS Survey provides new comprehensive empirical evidence on an educational mismatch in Italy. The main elements of the novelty, as reported in this paper derive, above all, from the very nature of the dataset: the adoption of Cedefop definitions allows us to provide more accurate measurements of the different aspects of the educational mismatch, while ensuring comparability across countries, although an international comparison goes beyond the aims of this paper. Any reader may compare our indicators of the educational mismatch with those computed relative to other EU countries. Moreover, the accuracy of Inapp-PLUS indicators is further increased by the exclusive use of direct interviews, as compared to indirect interviews (e.g., of a family member), which are common to other similar surveys. As discussed at more length in previous descriptions of the Inapp-PLUS

data bank (Mandrone and Radicchia 2012), direct interviews lead to the reduction of several cases of measurement errors when entering the details of educational, vocational training, and working experiences that prove to be ever more complex and, hence, not fully understood even by own family members.

This paper reports findings relative to three subjective indicators (validity of educational qualification, overeducation, and overskilling) and two objective indicators (one ISCO-based and one based on the modal value of a given profession). The indicators are calculated both for upper secondary school graduates and for university graduates. Moreover, a new synthetic indicator is proposed, which is constructed considering all other aforementioned indicators, in an attempt to reduce their respective limitations and identify different cases and levels of educational mismatch: no mismatch, partial mismatch, or full mismatch, depending on whether none, some or all of the five indicators suggest the presence of a mismatch. The described analysis points out that: subjective indicators are more coherent among themselves and cover all professions, even at the level of senior positions, in this way increasing the scope of observability of the phenomenon; statistical (so-called objective) indicators typically show higher values for overeducation, with a tendency to overestimate the phenomenon; more accurate measurements, with a higher grade of disaggregation per professional level, allow for more realistic results.

The synthetic indicator returns an absolute level of overeducation: it has a unanimous validation by all the indicators, always lower than 10% regardless of the specific group of individuals considered. This confirms the weak coherence among indicators. However, the highest value of the synthetic index, denoting the condition of overeducation according to all the indicators, is more frequent among individuals coming from poorer households (7%), which have found a job through the Public Employment Office (8%), a Temporary agency (7%) or through informal networks of friends and relatives (7%). In a way, this figure may be considered a bottom-line measure of overeducation that is coherent with all the definitions. Again, in presence of partial overeducation, the values are quite high, much higher than those from the indicators considered individually, which suggests the presence of widespread overeducation, over and beyond what individual indicators are capturing.

The inferential analysis identified the main determinants of overeducation, both from the supply and



demand side. From the perspective of the qualifications' supply side, overeducation decreases with age, coherently with the career mobility theory, and never disappears (in contrast to the same theory) being present also in advanced age, when one would expect that a good match has finally been obtained through labour mobility. This seems to confirm that overeducation may be a trap for those not managing to find a job in line with their level of competencies, which is particularly true for overskilling. Men are more frequently overeducated and so are those workers with lower skills in terms of foreign language and ICT knowledge. Overeducation is more frequent among people coming from lower-income families and with a humanistic educational background, as well as among those with some scientific educational background (geology and biology) that are less connected to access to a specific profession. In addition, the channel used to enter employment plays a role: the formal channel (open competitive selection, school, university, training institutes, or recruiting companies) better protect against overeducation than the informal network of family members, friends, and acquaintances, as well as self-proposals. Public and private employment agencies are also associated with a high level of overeducation.

From the demand side, we find the existence of a more common overeducation profile with indicators that are quite unanimous for private and non-profit companies rather than for public companies, for smaller rather than larger companies, in the agricultural and manufacturing sectors rather than the service and construction sectors. Overall, the analysis confirms the presence of important determinants not only from the

supply side, as often underlined in most contributions on the matter, but also from the demand side.

As for the effects, the analysis shows that overeducation is associated with a lower return. The wage penalty is higher for upper secondary graduates, with regression coefficients of 0.11 while the maximum wage penalty for tertiary graduates is expressed by a regression coefficient of 0.088. The double condition of overskilling and overeducation determines a wage penalty even of -0.17 for upper secondary graduates while for tertiary education it reaches only -0.13. This finding is confirmed with and without control for sample selection bias, although the evidence in favour of sample selection bias is circumstantial most likely because instrumental variables seem to be unable to fully account for sample selection.

These results may help education policymakers at all levels, as well as legislators, to identify more accurate measures to deal with the distortions of the national education system. From the supply side, an enhanced educational offer from schools and universities would be certainly desirable, which points to the development of the quality of human capital essentially through the expansion of work-related skills. Furthermore, from the demand side, incentives for private companies to hire graduates should be developed, and more resources should be allocated to R&D. University spin-offs and research centres should also be strengthened, and the creation of technology parks for the development of innovative ideas by graduate individuals should be favoured. All the sectors in the public administration need to contribute to this development, alongside private companies.

## References

- Attanasio O.P., Kaufmann K.M. (2017), Education Choices and Returns on the Labor and Marriage Markets. Evidence from Data on Subjective Expectations, *Journal of Economic Behavior & Organization*, 140, pp.35-55
- Buchel F., Van Ham M. (2003), Overeducation, regional labor markets, and spatial flexibility, *Journal of Urban Economics*, 53, n.3, pp.482-493
- Cappelli P.H. (2015), Skill gaps, skill shortages, and skill mismatches. Evidence and arguments for the United States, *ILR Review*, n.68, pp.251-290
- Capsada-Munsech Q. (2015), The Role of Social Origin and Field of Study on Graduates Overeducation. The Case of Italy, *Higher Education*, 69, n.50, pp.779-807
- Caroleo F.E., Pastore F. (2018), Overeducation at a Glance: Determinants and Wage Effects of the Educational Mismatch, Looking at the AlmaLaurea Data, *Social Indicators Research*, 137, n.3, pp.999-1032
- Caroleo F.E., Pastore F. (2016), Overeducation. A Disease of the School-to-Work Transition System, in Coppola G.L., O'Higgins N. (eds.), *Youth and their Future. Unemployment, Education and Health in Europe*, London and New York, Routledge, pp.36-56
- Caroleo F.E., Pastore F. (2013), L'overeducation in Italia. Le determinanti e gli effetti salariali nei dati AlmaLaurea, *Scuola democratica*, 2, n.2, pp.353-378
- Cedefop (2010), *Skill mismatch in Europe Europe's challenge is not just to improve skill levels, but to match people with the right skills to the right jobs*, Briefing note n.9023, Thessaloniki, Cedefop <<https://bit.ly/3ZvR48p>>

- Corsetti G., Mandrone E. (2012), Isfol PLUS Survey, in Mandrone E. (ed.) *Labour Economics: PLUS Empirical Studies*, Temi&Ricerche n.3, Cava de' Tirreni, Ediguida, pp.237-256
- Corsetti G., Mandrone E., Spizzichino A. (2014), L'Indagine PLUS, *Rivista Italiana di Economia Demografia e Statistica*, LXVIII, n.3-4, pp.1-8
- Croce G., Ghignoni E. (2016), Educational Mismatch and Spatial Flexibility in Italian Labour Market, *Education Economics*, 23, n.1, pp.25-46
- Croce G., Ghignoni E. (2012), Supply and Demand of Skilled Labour and Overeducation in Europe. A Country Level Analysis, *Comparative Economic Studies*, n.54, pp.413-439
- Cutillo A., Di Pietro G. (2006), The effects of overeducation on wages in Italy: a bivariate selectivity approach, *International Journal of Manpower*, 27, n.2, pp.143-168
- Davia M.A., McGuinness S., O'Connell P.J. (2017), Determinants of regional differences in rates of overeducation in Europe, *Social Science Research*, 63, pp.67-80
- Dolton P., Vignoles A. (2000), The incidence and effects of overeducation in the UK graduate labour market, *Economics of Education Review*, 19, n.2, pp.179-198
- Flisi S., Goglio V., Meroni E.C., Rodrigues M., Vera-Toscano E. (2017), Measuring occupational mismatch: overeducation and overskill in Europe. Evidence from PIAAC, *Social Indicators Research*, 131, n.3, pp.1211-1249
- Franzini M., Raitano M. (2012), Few and underutilized? Overeducation of Italian graduates, in Mandrone E. (ed.) *Labour Economics: PLUS Empirical Studies*, Temi&Ricerche n.3, Cava de' Tirreni, Ediguida, pp.17-52
- Giammatteo M. (2009), L'indagine campionaria Isfol-PLUS: contenuti metodologici e implementazione, Studi Isfol n.3, Roma, Isfol
- Hartog J. (2000), Overeducation and earnings: where are we, where should we go?, *Economics of education review*, 19, n.2, pp.131-147
- Heckman J.J. (1979), Sample selection bias as a specification error, *Econometrica*, 47, n.1, pp.153-62
- Hungerford T., Solon G. (1987), Sheepskin effects in the returns to education, *The Review of Economics and Statistics*, 69, n.1, pp.175-177
- Istat (2012), *Generazioni a confronto: come cambiano i percorsi verso la vita adulta*, Roma, Istat
- Johnes G. (2019), The Incidence of and Returns to 'Overeducation': PIAAC Evidence on the G7, *Minerva*, n.57, pp.85-107
- Kleibrink J. (2016), Inept or Badly Matched? Effects of Educational Mismatch in the Labor Market, *Labour*, 30, n.1, pp.88-108
- Korpi T., Tåhlin M. (2009), Educational Mismatch, Wages, and Wage Growth. Overeducation in Sweden, 1974-2000, *Labour Economics*, 16, n.2, pp.183-193
- Kucel A. (2011), Literature Survey of the Incidence of Overeducation. A Sociological Approach, *Revista Espanola de Investigaciones Sociológicas (REIS)*, n.134, pp.125-142
- Leuven E., Oosterbeek H. (2011), Overeducation and Mismatch in the Labour Market, in Hanushek E.A., Machin S., Woessmann L. (eds.), *Handbook of the Economics of Education*, pp.283-326
- Mandrone E. (2012), *Labour Economics: Plus empirical studies*, Temi&Ricerche n.3, Cava de' Tirreni, Ediguida
- Mandrone E. (2008), *La riclassificazione del lavoro tra occupazione standard e atipica: l'indagine Isfol Plus 2006*, Studi Isfol n.1, Roma, Isfol
- Mandrone E., Radicchia D. (2012), *Indagine PLUS. Il mondo del lavoro tra forma e sostanza. Terza annualità*, I libri del Fondo sociale europeo n. 167, Roma, Isfol
- Matkovic T., Kogan I. (2012), All or Nothing? The Consequences of Tertiary Education Non-Completion in Croatia and Serbia, *European Sociological Review*, 28, n.6, Special Issue, pp.755-770
- McGuinness S., Sloane P.J. (2010), «Educational/skill mismatch» nel mercato del lavoro dei laureati: un'analisi comparata, in AlmaLaurea, *XII Rapporto sulla condizione occupazionale dei laureati. Investimenti in capitale umano nel future di Italia ed Europa*, Bologna, il Mulino, pp.101-115
- Meliciani V., Radicchia D. (2015), *Overskill and overeducation in the Italian labour market: the role of fields of study*, presented at *Birkbeck CIMR Workshop: Risky Skills*, 10 luglio
- Meliciani V., Radicchia D. (2009), *L'intermediazione informale in Italia: è vantaggioso risparmiare sui costi di ricerca di un lavoro?*, Studi Isfol n.5, Roma, Isfol
- Montt G. (2017), Field-of-study mismatch and overqualification: labour market correlates and their wage penalty, *IZA Journal of Labor Economics*, 6, n.1, p.2
- Nicaise I. (2001), Human capital, reservation wages and job competition: Heckman's lambda reinterpreted, *Applied Economics*, 33, n.3, pp.309-15
- OECD (2015), *Education at glance*, Paris, OECD Publishing
- Ortiz L., Kucel A. (2008), Do fields of study matter for overeducation: the cases of Spain and Germany, *International Journal of Comparative Sociology*, 49, n.4-5, pp.305-327
- Pouliakas K. (2012), The skill mismatch challenge in Europe, in *Employment and social developments in Europe*, Brussels, Council of the European Union <<https://bit.ly/3tdTw4D>>
- Quintini G. (2011), *Over-Qualified or Under-Skilled: A Review of Existing Literature*, OECD Social, Employment and Migration working paper n.121 <<https://bit.ly/3DLy6k4>>
- Rossen A., Boll C., Wolf A. (2019), Patterns of Overeducation in Europe: The Role of Field of Study, *IZA Journal of Labor Policy*, 9, n.1 <DOI: <https://doi.org/10.2478/izajolp-2019-0003>>

- Sicherman N., Galor O. (1990), A theory of career mobility, *Journal of Political Economy*, 98, n.1, pp.169-192
- Silles M. (2008), Sheepskin Effects in the Returns to Education, February, *Applied Economics Letters*, 15, n.3, pp.217-219
- Sloane P. (2003), Much Ado about Nothing? What does the Overeducation Literature really Tells us?, in Büchel, F., de Grip A., Mertens A. (eds.), *Overeducation in Europe: Current Issues in Theory and Practice*, Cheltenham, Edward Elgar Publishing, pp.11-48
- Sloane P.J., Battu H., Seaman P.T. (1999), Overeducation, undereducation and the British labour market, *Applied Economics*, 31, n.11, pp.1437-1453
- Thurow L. (1979), A Job-Competition Model, in Piore M. (ed.), *Unemployment and Inflation: Institutional and Structuralist Views*, New York, Routledge, pp.17-32

### Emiliano Mandrone

e.mandrone@inapp.org

Economist, First Researcher at Istat, currently on secondment at Inapp. He has taught at the University of Turin, Urbino and Rome. He has designed and conducted the Inapp-PLUS survey for over fifteen years. He coordinates the Inapp project *Revisione degli indicatori socioeconomici alla luce della transizione digitale*. He collaborates with *lavoce.info*, *agendadigitale.eu* and *ilfattoquotidiano.it*. Recent publications include: Digital oddities: technological change and cultural elaboration, *Sinapsi*, 2021; Reverse Urbanization How Remote Working and Technology are Changing Cities, *Global Journal of Human-Social Science*, 2021.

### Debora Radicchia

debora.radicchia@anpal.gov.it

Researcher at the Office for Statistics and Methodological Support of Anpal. She was co-responsible for the PLUS Survey from 2004 to 2016. Recent publications include: Informal networks, spatial mobility and overeducation in the Italian labour market, with V. Meliciani, *The Annals of Regional Science*, 2016; *Disoccupazione amministrativa: prima analisi delle DID*, with Baronio G., Linfante G., Mattei L., Porcari S., Rosato M., Focus Anpal. Metodologie e Approfondimenti, 2021.

### Antonella Rocca

antonella.rocca@uniparthenope.it

Associate Professor of Economic Statistics, University of Naples Parthenope. She teaches Economic Statistics in degree and doctorate courses. She collaborates as an expert on the labour market with the European Commission and is the representative for Italy of the COST Action 18213 Rural Neet Youth Network. Recent publications include: Where to Start from? Job-finding Intentions among young adults from 11 European countries, with Tosun J. and Simoes F., *Social Indicators Research*, 2022; Some young people have all the luck! The duration dependence of the school-to-work transition in Europe, *Labour Economics*, with Pastore F., Quintano C., 2021.

### Francesco Pastore

Associate Professor of Political Economy at the University of Campania Luigi Vanvitelli, Research fellow IZA; Country leader for Italy and leader of the cluster on *school-to-work transition* of the Global Labor Organization; Vice-President of the European Association of Comparative Economic Studies. Recent publications include: The Duration of the School-to-Work transition in Italy and in other European countries: A flexible baseline hazard interpretation, *International Journal of Manpower*, with Quintano C., Rocca A., 2022; Some young people have all the luck! The duration dependence of the school-to-work transition in Europe, *Labour Economics*, with Quintano C., Rocca A., 2021.

### Claudio Quintano

claudio.quintano@emerito.uniparthenope.it

Professor Emeritus, former Full Professor of Economic Statistics at the University of Naples Parthenope. He teaches, at the University of Naples Suor Orsola Benincasa, 'The measures of tomorrow to achieve the UN sustainability goals'. He has been a member of the Presidential Council of Istat, and a member of the Commission for the Guarantee of Statistical Information of the Presidency of the Council of Ministers. Since March 2020 he has collaborated with *Sussidiario.net* and *Corriere del Mezzogiorno-Corriere della Sera*. Recent publications include: Stuck at a crossroads? The duration of the Italian school-to-work transition, *International Journal of Manpower*, with Pastore F., Rocca A., 2020; The Duration of the School-to-Work transition in Italy and in other European countries: A flexible baseline hazard interpretation, *International Journal of Manpower*, with Pastore F., Rocca A., 2022.

Due to exceptional circumstances, the present paper hasn't undergone any peer-review.