

SENZA WELFARE?
Federalismo e diritti di cittadinanza nel modello mediterraneo
Napoli, 30 settembre – 2 ottobre 2010

Overeducation And Spatial Flexibility In Italian Local Labour Markets

Giuseppe Croce * - Emanuela Ghignoni **

Paper presentato alla Terza Conferenza annuale ESPAnet Italia 2010

SESSIONE: NR.3C - LA "GEOGRAFIA" DELLE DISUGUAGLIANZE ECONOMICHE IN ITALIA: OLTRE I
DATI MONETARI

* Sapienza – University of Rome, Department of Economics and Law. Via Castro Laurenziano, 9, 00161, Roma; giuseppe.croce@uniroma1.it.

** Sapienza – University of Rome, Department of Economics and Law. Via Castro Laurenziano, 9, 00161, Roma; Emanuela.ghignoni@uniroma1.it.

OVEREDUCATION AND SPATIAL FLEXIBILITY IN ITALIAN LOCAL LABOUR MARKETS

Giuseppe Croce^a, Emanuela Ghignoni^b

VERY PRELIMINAR VERSION – DO NOT QUOTE

Abstract

Keywords: Overeducation, Spatial flexibility, Local labour markets

JEL classification: J21, J61, J62

1. Introduction

Traditional labour market theory often state that individuals search for work in the global labour market, whilst most workers look for jobs in the local one. In this framework, the spatial mismatch hypothesis explains the presence of unemployment with the lack of local job opportunity in combination with limited spatial flexibility of the workforce (Preston and McLafferty, 1999).

In order to widen this line of research, an increasing literature points out the need of considering a broader concept of labour market disadvantage which accounts not only for those people who are unemployed, but also for individuals who are underemployed, include therein overeducated workers (Hensen and de Vries, 2004; Baum et al., 2008), and recognizes that the causes of underemployment are not only the skill mismatch at national level, but also the spatial mismatch between where workers live and where jobs are located (Van Ham, 2002).

Nevertheless, the spatial effects of underemployment, and in particular of overeducation, have been neglected for long, with the commendable exception of the theory of differential overeducation, about the job seeking behaviour of married women (Frank, 1978; Büchel and Battu, 2003).

In this framework, the seminal paper of Büchel and van Ham (2003) highlighted the role of regional labour markets characteristics as a potential explanatory variable of global overeducation and not only of the differential overeducation of married women.

According to previous results coming from this strand of literature, we estimate a model where workers' spatial mobility is considered as a factor affecting the probability of overeducation. Indeed our dataset allows us to detect both commuters and movers for work reasons. We also include in the model other variables characterising the local labour markets where the worker resides in order to assess their impact on the probability of overeducation.

The paper is organized as follows. Section 2 contains a literature review. Section 3 describes methodology and data and Section 4 contains the main results of econometric estimates. In Section 5 we discuss some concluding remarks and policy implications.

2. Literature review

Overeducation represents a multifaceted phenomenon in the labour market of advanced economies as it can be caused by a number of factors and is consistent with different theories (Hartog 2000, Sloane 2003). It can derive from an excess of highly educated labour supply, when

^a Sapienza – University of Rome, Department of Economics and Law.

^b Sapienza – University of Rome, Department of Economics and Law.

the aggregate amount of workers with a tertiary degree or a high school diploma exceeds the amount of available jobs requiring the corresponding level of education. If wages are not enough flexible and firms do not adapt their jobs to workers' characteristics, such mismatch tends to persist. Possible explanations in this line are given by the job competition and job assignment models (Thurow 1975, Sattinger 1993).

Overeducation can be reconciled even with human capital theory if the supply excess is seen as a transitory state, or if individual unobserved ability is considered or, finally, by assuming substitutability between education and on-the-job training (McGuinness 2006).

Other explanations do not presuppose aggregate demand-supply imbalances and point to labour market imperfections as possible causes of overeducation. High search or mobility costs, e.g. resulting from information imperfections and family ties, can shape the distribution of the risk of being overeducated among heterogeneous workers. Following Frank (1978), McGoldrick and Robst (1996), and Buchel and van Ham (2002), the spatial flexibility hypothesis focuses, in particular, on spatially based explanations of overeducation looking for factors affecting the individual propensity/ability to commute or to migrate.

The central aspect of the analysis by Büchel and van Ham (2003) consists in analyzing which is the role of job opportunities in local labour markets (measured by regional unemployment rates) and commuting possibilities (measured by the travelling time to the nearest job agglomeration and the availability of a car for personal use) to explain the probability of being overeducated. As in Frank (1978), geographical restrictions play a key role in explaining overeducation, but Büchel and Van Ham extend the effect to all workers (and not only women) and they consider the possibility of extending the job search to other local labour markets through commuting. Their results for the German labour market show that regional variables related with the spatial distribution of employments explain overeducation. In particular, a higher mobility of individuals (owning a car) permits to increase the "effective" size of the labour market, which decreases the probability of being overeducated. By contrast, local unemployment rate, which should be an index of competition for jobs in the local labour market, have a significant (negative) impact on the probability of being employed but does not affect the probability of being overeducated.

In this line, Sanroma and Ramos (2004) analyse the possible effects on the probability of overeducation of some variables related to the territory in Spanish local labour markets. In particular, they proxied the local labour market size by including in the model a series of dummies indicating the number of inhabitants in the town of residence of the individual, and take into account the possibility of widening the spatial job search through the inclusion in the model of two variable: the availability of private transport (at individual level) and the number of road kilometres respect to the number of cars (at provincial level). Their results, in line with those obtained for the German labour market by Büchel and van Ham (2003), show that living in a small town (that is, having less local job opportunities) increases the probability of being overeducated, whilst the possibility of increasing local labour market's size through searching job in more distant labour markets, thanks to the availability of private transport and a good road network, significantly reduces overeducation.

An interesting study on French data (Cahuzac and di Paola, 2004) put in evidence the importance of employment density as a measure of local labour market's size. The authors use a series of dummies based on the notion of "poles urbains", (that is, municipalities which offer more than 5,000 jobs) under the hypothesis that employment density could have two opposing effects on the probability of overeducation: on the one hand, on denser labour markets the numerous job offers can increase the probability of finding a (well-matched) job but, on the other hand, denser labour markets could decrease the chance of finding a job matching workers' skill due to competition among them. The authors take also into account the demographic evolution of local population and the local unemployment rate, as a measure of competitions for jobs. As to the commuting possibility they include in the model the average travelling time residence-work at local level, which should account for the quality of local transport infrastructures, as well as the average

distance between places of residence and job agglomerations. Results show that overeducation is lower in denser and more populated areas and in local labour markets characterised by lower unemployment rates. This finding contradicts the outcome of Büchel and van Ham (2003), in which regional unemployment rates do not affect overeducation. Actually, this discrepancy could depend on the different definition of overeducation used in the different analysis: a subjective one in Büchel and van Ham and a statistical one in Cahuzac and di Paola. Finally, overeducation increases in those local labour market in which the average travelling time residence-work is longer. Thus, an easier access to a larger concentration of jobs seems to lead to better matches.

An analysis of overeducation in the Australian labour markets (Linsley, 2005) assumes that the probability of overeducation is strongly affected by spatial constraints, which are greater for married women with children and for individuals without access to adequate transport options or who live long distances from urban labour markets. In practise, the access to a motor vehicle is used to capture an individual's capacity to commute, and some location variables (major city, inner regional areas, outer regional and remote areas) to capture the size of local labour market and the distance between individual's residence and agglomerations of jobs. The analysis concludes that women (married or not) living in outer regional or remote areas have a significantly higher incidence of overeducation than their counterparts in inner regional areas and major cities. By contrast, access to a motor vehicle has no discernible effect on the probability of overeducation. This apparently contradicts the spatial mobility theory. However, as already highlighted by Büchel and van Ham (2003), this result may be due to endogeneity. If overeducation is negatively correlated with earnings, then the overeducated are less likely to be able to afford a motor vehicle.

Some interesting recent analysis of overeducation on Italian data¹ pointed out the existence of high levels of persistent overeducation, but failed to take into account the impact of the local labour markets characteristics, although territorial dimension has proved to be a very important variable in determining individuals' outcomes in Italian labour markets².

3. Methodology and data

Overeducation describes the extent to which an individual possesses a level of education in excess of that which is required for his/her particular job. Economic literature suggests five alternative measures of overeducation:

1. The *objective* measure: Professional job analysts try to specify the required level and type of education in a particular occupation and, through this, provide a classification of occupation into a requirement ranking (Rumberger, 1987).
2. The (*direct* or *indirect*) *subjective* measure (Duncan and Hoffman, 1981; Hartog and Oosterbeek, 1988: This approach is based on worker's self-assessment about his/her job-skill match. The worker can be directly asked "Is your educational degree necessary to get or to do your job?", (or "Are the competencies you acquired during schooling education fully utilised in your job?"). In alternative, the worker can be indirectly asked: "In your opinion, which is the more adequate level of education required to carry out your job?". In this case overeducation is identified by comparing worker's answer with his/her level of education.
3. The *empirical* (or *statistical*) measure: This approach uses the distribution of schooling years in a given occupation or a group of occupations. Most commonly, individuals are defined to be overeducated if their schooling level is more than one standard deviation above the mean (Verdugo and Verdugo, 1988) or mode (Kiker et al., 1997) of all individuals in that occupation.

¹. See Franzini and Raitano, 2009; Ordine and Rose, 2009a, 2009b.

². See Dalmazzo and De Blasio, 2007; Brunello and De Paola, 2008; Di Addario and Patacchini, 2008; Bratti and Leombruni, 2009; Croce and Ghignoni, 2009.

4. The *income ratio* measure: The individual potential income is estimated with a stochastic earning frontier. Then, overeducation is measured as the income ratio between potential income and actual income (Jensen, 2003).
5. The *competences' frontier* method: This method is based on the estimation of an indifference curve on the plane “years of education/years of experience” for each occupation, characterised by a negative slope illustrating the existence of a substitutability relation (at macroeconomic level) between these two basic elements of individuals' competences (Ghignoni, 2001).

Unfortunately, these methods provide quite different results when applied to the same database, and the application of one or other measurement is generally determined by the available information.

In this framework, Isfol-Plus survey is considered a very useful dataset in order to analyse overeducation in Italy (Franzini and Raitano, 2009). In this survey overeducation is assessed on the grounds of a (direct) subjective method (question v430_1: *Is your educational degree necessary to perform your job?*). In any case, it worth noticing that interviewed workers do not specify if they refer to a formal or substantial necessity. Moreover they do not indicate the effective degree of utilisation of the competences acquired during their educational path, and it is not possible to distinguish genuine cases of overeducation (a job below the individual educational level) from mismatches due to having a job outside the individual educational field. Actually, these caveats are common to all overeducation analysis based on subjective measures. However, Isfol-Plus survey provides very detailed information on schooling levels and occupations, useful to built statistical measures of overeducation. Moreover, the abundance of information and the sample numerosness allow controlling for a variety of explanatory variables, some of which can be considered as proxy for individual ability (school grades, regular schooling paths) and practical skills (in particular, computer and language skills).

To carry out our analysis we exploit individual data drawn from the Isfol-Plus 2005 and data from various sources on LLMs. We select individual with an upper secondary or a tertiary degree (26,323 individuals), aged 20 or (25,522 individuals), who declare to be employed, unemployed or housewife, which leads to a further sample reduction (19,335 individuals).

Table 1 presents some descriptive statistics of the variables used in the empirical analysis.

The objective of the paper is to identify the explanatory factors of the probability of being overeducated, keeping into account not only individual and firm-specific characteristics, but also some variables related to local labour markets (LLMs). In this case, the following probit specification:

$$Prob\{OVER_{ij} = 1\} = \Phi\{\beta X_{ij} + \delta Y_j + \varepsilon_{ij}\} \quad [1]$$

permits to estimate the variation in the probability of being overeducated after marginal variations of the explanatory variables.

In this case OVER is a dichotomous variable which assume value 1 if the individual is overeducated and value 0 otherwise. X is a vector of individual and firm-specific observable characteristics, Y is a vector of area-specific effects and the indices i and j refer to the individual and the local area.

For the purpose of this paper local areas correspond to LLMs which can be defined as travel-to-work areas corresponding to groups of neighbouring municipalities, belonging or not to the same Region, aggregated on the basis of data on daily commuting for working reasons in a way that most of the residents also work in the area (Istat, 1997).

As pointed out by Moulton (1990), some problems could emerge in regressions using both aggregate data and data on micro units' characteristics as explanatory variables, due to the possibility that the random disturbances in the regression are correlated within groups. To handle

with this problem we calculated robust standard errors corrected for the potential clustering of the residuals at the LLM level.

An additional aspect that should be mentioned before starting the empirical analysis is that overeducation may concern only employed workers and we can observe the dependent variable of equation [1] only if the individual actually works. As well-known, analysing overeducation while restricting the sample to the employed could lead to biased results³. Indeed, if no suitable employment is available, unemployment can be chosen as a strategy to avoid overeducation. In this case the most likely to be overeducated would be those least likely to enter employment.

To deal with this problem we estimated the following bivariate probit model with sample selection⁴:

$$\begin{aligned} P_{OVER} &= \alpha + \beta X + \delta Y + \varepsilon_1 \\ P_{WORK} &= bZ + \varepsilon_2 \end{aligned} \quad [2]$$

where:

$$P_{OVER} = \int_0^1 \text{if the individual is overeducated} \quad \text{and} \quad P_{WORK} = \int_0^1 \text{if the individual actually works} \\ \int_0 \text{otherwise} \quad \int_0 \text{otherwise}$$

X and Y are, respectively, a vector of individual and firm-specific characteristics, and a vector of local labour market characteristics influencing the probability of overeducation, whilst Z is a set of variables that influence the probability of working.

The key element of this strategy is to select at least one instrument variable that affects only the access in employment, but not the overeducation risk. We decide to use as selection variable a dummy indicating those individuals who pay a rent or a mortgage loan for accommodation. Thus, we hypothesize that the presence of these financial burdens is strongly (positively) correlated with the probability of being employed. However, this variable, conditional on working, should not have a direct influence on labour market outcomes once we control for individual characteristics and job and firm-specific variables.

4. Estimations results

Our baseline estimation (see table 2) include:

- individual variables: age, experience, level of education, type of contracts, municipality size, a series of cross dummies gender/marital status/location (useful to control for the presence of differential overeducation (Büchel and Battu, 2003), the presence of children aged 0-12, schooling ability (school grades and “regularity” of the educational path), linguistic/informatics skills and father education. We also include the daily commuting time and a dummy for movers;
- firm specific variables: branch of economic activity and firm size;
- local variables: the percentage of upper secondary and tertiary graduates, the average firm size in the LLM, the employment density, the local rate of unemployment, a dummy for the industrial districts and some variables indicating the quality of transport infrastructures at local level (the percentage of mountain area in the LLM and the number of accidents per 1000 cars).

As in Büchel and Van Ham (2003) age seems not to influence overeducation, while permanent and full time employees are less overeducated than fixed term and part time workers.

³ . See Ordine and Rose, 2009; Büchel and Van Ham, 2003.

⁴ . Van de Ven and Van Praag, 1981.

The cross dummies gender/marital status/location are not significant. This series of dummies should have control the differential overeducation theory. According to this theory married women in rural areas should suffer higher risks of overeducation than other workers. As in Büchel and Battu, 2003, the theory is not verified.

Both schooling and job-related skills appear to reduce the risks of overeducation, as well as being Italian.

In accordance with our hypothesis, individual who accept longer daily commuting time and movers for working reasons bear minor overeducation risks. Thus, individual mobility is recompensed with a better skills/job match.

As to local variables, overeducation is lower in LLMs with higher employment density and in LLMs who are industrial districts. By contrast, overeducation is higher in LLM who are university centre. As in Büchel and Van Ham (2003), local unemployment rate does not have a significant impact on overeducation.

As ascertained by previous literature, being in employment and overeducated cannot be seen a priori as independent statuses in the labour market since factors affecting the probability of employment could also impinge on overeducation risk. For this reason results from the simple probit model could be unreliable and a probit model with sample selection has to be estimated.

The estimate of employment equation (Table 3.a) reveals that in our sample the probability of being employed decreases with age and is lower for university graduates than for high school graduates. Moreover married women tend to be less employed, and married men more employed regardless of the kind of their area of residence (city or rural area). The probability of employment is lower for people with children. Among the variables referring to the local labour market, only the unemployment rate coefficient, as expected, is negative and significant.

It must be remarked that, as it was assumed, the loan/rent dummy is highly statistically significant: people paying a loan/rent are more likely to be in employment.

Turning to the overeducation equation, our findings (Table 3.b) show that a quite limited number of individual and job characteristics affects the probability of being overeducated along with a few variables representing local labour market features. Most important, both the variables describing mobility behaviour proved to be relevant.

As for the individual characteristics, age, as well as education (holding a degree rather than a high school diploma) are uninflential while the specific experience has a negative effect and is almost statistically significant. In line with Hartog (2000) this result, according with career mobility theories (Linsley, 2005a, 2005b), suggests that workers who stay longer in their job are those in a more qualified position, facing a lower probability of overeducation. Unlike results for the employment equation, having children is not relevant for overeducation. Foreigners are more likely to be overeducated than Italians. It is worth noticing that interactions terms between the dummies for city/rural area, gender and married/single are never significant despite predictions of the hypothesis of differential overeducation (Frank, 1978; Büchel and Battu, 2003).

Individual ability, as proxied by the marks at the end of school or university and by the regularity of school or university path, results to reduce the risk of overeducation. The same negative and statistically significant impact is associated to the possession of very general competences as the ability of surfing the Internet and speaking English. Turning to job characteristics, the employment contract results to be an influential factor. Indeed both a permanent and a full time contract reduce the likelihood of overeducation in comparison to, respectively, temporary and part-time contracts. This outcome rules out that the acceptance of an atypical contract can be traded-off for a well-matched position. On the contrary, it suggests that temporary and part-time employment characterizes the secondary sector in the labour market where also the risk of an unsatisfactory match between education possessed and required is higher. Also workers employed in medium sized firms (50-499 employees) or in Distribution and Personal Services face a higher risk of overeducation.

As for the territorial aspects, the dummy for the macroarea Centre has a positive and highly significant effect, pointing out that in this area there could be an aggregate imbalance in the labour market between the composition of demand and supply which makes more likely overeducation. On the other hand, the share of people holding a tertiary or a high school degree in the local labour market is far from being significant. This outcome does not confirm the hypothesis that the relative weight of highly educated in the labour supply at the local level is associated with overeducation. Moreover neither the size of the municipality nor the local unemployment rate is relevant, while to live in denser areas or in local labour markets corresponding to an industrial district reduces the probability of being overeducated. Thus, a higher employment density enlarges the possibilities of choice and improves the outcome of the matching process. At the same time, it is likely that demand and supply of skills tend to be more homogeneous in the industrial districts.

Finally, our key-result is that both the commuting time and having moved for working reasons related to the current job are significant and show the expected sign. The risk of overeducation decreases with time spent commuting, moreover movers are likely to be less overeducated. This evidence is consistent with the hypothesis of spatial flexibility which predicts that workers able to search for job in a larger area improve their chances of fully valorize their human capital. More in general, it points out that frictions and barriers in the labour market give a possible explanation of overeducation, which could be independent of aggregate imbalances between the composition of demand and supply of skills.

Note that the coefficient on the selection term (ρ) is not statistically different from zero (see table 3.a). This result is confirmed by the Wald test for independent equations and suggest that ignoring selection into work should not introduce bias into the overeducation probabilities.

5. Concluding remarks and policy implications

Among the various theories which have been advanced to explain overeducation, this paper goal was to test the relevance of the so called spatial flexibility hypothesis. According to it, overeducation arises when workers cannot afford the costs of mobility needed to search for a suitable job if demand and supply of skills are not homogeneously distributed on a spatial basis. It is worth noticing that this explanation of overeducation points to frictions and barriers determining a less than perfect mobility in the labour market rather than to aggregate imbalances between demand and supply of skills or to a low quality of learning due to the services provided by the educational system.

Our dataset allowed us to consider the commuting time and if the worker have moved for working reasons related to her/his current job, which represent two fundamental dimensions of the spatial flexibility. Besides these two factors, we included in the model a set of variables characterising the local labour market where the worker resides.

To overcome the problem of selection bias due to the correlation between employment and overeducation we estimated a bivariate probit model with sample selection. The selection variable we added to the employment equation was a dummy detecting individuals who have a loan/rent to pay for. As it was assumed, this variable proved to be highly relevant in affecting the employment probability. Moreover, according to our results the two equations could be considered as not correlated.

Our most important finding is that spatial flexibility matters. Indeed the risk of overeducation decreases with commuting time and *volunteer* movers are less overeducated than stayers. Even though this result does not rule out that also demand-supply imbalances can be relevant, it implies that friction and barriers limiting spatial mobility worsen the matching between required and possessed education in the labour market.

As for policy implications, such evidence suggests that a more efficient valorisation of human capital could be reached by reducing both monetary and non monetary costs of

mobility. This would require flexibility of working time, the availability of services for families, like child care, and, of course, proper transportation facilities and services. Moreover, the conditions of the housing market represent an important factor to boost mobility. Finally, even the spread of teleworking can be regarded as a technological and organisational innovation which could be helpful in reducing the incidence of overeducation in a number of branches of activities and occupations.

Annex 1 – Tables

Table 1 – Descriptive statistics

Variable description	Employed			Full sample		
	N. Obs	Mean	Std. Dev.	N. Obs	Mean	Std. Dev.
Dependent variable						
Overeducated = 1	12694	0.333	0.471			
Employed = 1				19335	0.657	0.475
Individual characteristics						
Age	12694	39.492	12.148	19335	37.726	11.466
Age square	12694	1707.22	993.028	19335	1554.71	920.759
Specific experience	12694	11.991	10.969	12694	11.991	10.969
Specific experience square	12694	264.077	371.881	12694	264.077	371.881
Generic experience	12694	17.456	12.000	12694	17.456	12.000
Generic experience square	12694	448.707	482.202	12694	448.707	482.202
<i>Upper secondary education</i>	12694	0.667	0.471	19335	0.711	0.453
Tertiary education	12694	0.312	0.463	19335	0.271	0.445
Post tertiary education	12694	0.022	0.146	19335	0.017	0.130
Permanent contract	12694	0.668	0.471	12694	0.668	0.471
Full time	10010	0.851	0.356	10010	0.851	0.356
<i>Less than 10,000 inhabitants</i>	12694	0.272	0.445	19335	0.284	0.451
10,001 – 100,000 inhabitants	12694	0.358	0.479	19335	0.375	0.484
More than 100,000 inhabitants	12694	0.371	0.483	19335	0.341	0.474
City woman married	12694	0.103	0.304	19335	0.105	0.306
City man married	12694	0.073	0.261	19335	0.052	0.223
Rural woman married	12694	0.206	0.404	19335	0.278	0.448
Rural man married	12694	0.182	0.386	19335	0.132	0.338
City woman single	12694	0.082	0.275	19335	0.075	0.263
<i>City man single</i>	12694	0.057	0.232	19335	0.053	0.224
Rural woman single	12694	0.144	0.351	19335	0.161	0.367
Rural man single	12694	0.153	0.360	19335	0.145	0.352
Children 0-12	12694	0.779	0.415	19335	0.816	0.387
High marks	12694	0.509	0.500	19335	0.479	0.500
Regular educational path	12694	0.777	0.416	19335	0.768	0.422
Internet	12694	0.854	0.353	19335	0.809	0.393
English	12694	0.354	0.478	19335	0.342	0.474
Daily commuting time (home to work)	10903	20.806	20.197	10903	20.806	20.197
Movers	12694	0.123	0.328	12694	0.123	0.328
<i>Father education: max lower secondary</i>	11985	0.678	0.467	18201	0.692	0.462
Father education: Upper secondary	11985	0.232	0.422	18201	0.225	0.418
Father education: Tertiary degree	11985	0.090	0.286	18201	0.083	0.276
Foreigner	12694	0.011	0.104	19335	0.014	0.118
North West	12694	0.240	0.427	19335	0.213	0.410
North East	12694	0.214	0.410	19335	0.189	0.391
Centre	12694	0.207	0.405	19335	0.203	0.402
South	12694	0.238	0.426	19335	0.276	0.447
Islands	12694	0.100	0.301	19335	0.118	0.323
Firms characteristics	12562	0.199	0.400	12562	0.199	0.400
Production of goods	12562	0.146	0.353	12562	0.146	0.353
Production services	12562	0.192	0.394	12562	0.192	0.394
Distribution services	12562	0.056	0.230	12562	0.056	0.230
Personal services	12562	0.407	0.491	12562	0.407	0.491
Social services	12562	0.199	0.400	12562	0.199	0.400
Small firm (10-49 employees)	7839	0.707	0.455	7839	0.707	0.455
Medium firm (50-499 employees)	7839	0.215	0.411	7839	0.215	0.411
Large firm (500 or more employees)	7839	0.077	0.267	7839	0.077	0.267
LLM characteristics						
Mountain area (km2)	12694	320.053	346.864	19335	307.900	345.640
University centre	12694	0.541	0.498	19335	0.516	0.500
% Foreigner	12694	4.665	2.621	19335	4.366	2.652
Average firm size	12694	8.005	2.898	19335	7.737	2.963
Industrial district	12694	0.180	0.384	19335	0.169	0.375
(log) Employment density	12694	4.811	1.251	19335	4.766	1.273
% Upper Secondary and Tertiary degree	12694	34.861	5.316	19335	34.460	5.480
Accident/1000 cars	12694	7.385	3.403	19335	7.100	3.405
Local unemployment rate	12694	5.319	3.706	19335	5.831	3.904
Selection variable						
<i>Rent/Loan for accommodation</i>	12694	0.189	0.392	19335	0.174	0.379

Table 2 – Probability of overeducation (without selection bias)

Variables	Coef.	Robust Std. Err.	z	P> z
Individual variables				
Age	0.02265	0.02196	1.03	0.302
Age square	-0.00030	0.00028	-1.07	0.285
Specific experience	-0.01761	0.01080	-1.63	0.103
Specific experience_square	0.00024	0.00029	0.82	0.410
Generic experience	-0.00126	0.01060	-0.12	0.905
Generic experience_square	0.00010	0.00027	0.35	0.723
<i>Education (ref. secondary ed.)</i>				
Tertiary education	-0.00910	0.07756	-0.12	0.907
Post tertiary education	-0.38078	0.29037	-1.31	0.190
Permanent contract	-0.10444	0.05177	-2.02	0.044
Full time	-0.22531	0.05749	-3.92	0.000
<i>Municipality size (ref. 0-10,000)</i>				
10,001 – 100,000	-0.02376	0.05548	-0.43	0.668
More than 100,000	-0.01564	0.07713	-0.20	0.839
<i>Gender/Marital status/location (ref. City man single)</i>				
City woman married	-0.12740	0.09687	-1.32	0.188
City man married	-0.08063	0.11660	-0.69	0.489
Rural woman married	-0.12896	0.10660	-1.21	0.226
Rural man married	-0.15041	0.10127	-1.49	0.137
City woman single	-0.00541	0.14424	-0.04	0.970
Rural woman single	-0.06199	0.10576	-0.59	0.558
Rural man single	-0.02724	0.10669	-0.26	0.798
Children 0-12	-0.05902	0.04963	-1.19	0.234
High marks	-0.23797	0.04034	-5.90	0.000
Regular educational path	-0.09000	0.03921	-2.30	0.022
Internet	-0.41066	0.05855	-7.01	0.000
English	-0.16905	0.04732	-3.57	0.000
Daily commuting time (home to work)	-0.00287	0.00101	-2.86	0.004
Movers	-0.20350	0.07820	-2.60	0.009
<i>Father education (max lower secondary)</i>				
Upper secondary	-0.04625	0.05203	-0.89	0.374
Tertiary degree	0.03794	0.07691	0.49	0.622
Foreigner	0.77063	0.14226	5.42	0.000
<i>Macroarea (North West)</i>				
North East	-0.08889	0.05723	-1.55	0.120
Centre	0.13569	0.05701	2.38	0.017
South	0.11013	0.09953	1.11	0.268
Islands	0.00080	0.13234	0.01	0.995
Firm characteristics				
<i>Firm size (small: 10- 49)</i>				
Medium (50-499)	0.14013	0.05094	2.75	0.006
Large (500 or more)	0.09812	0.07772	1.26	0.207
<i>Branch of economic activity (ref. production of goods)</i>				
Production services	-0.10310	0.05726	-1.80	0.072
Distribution services	0.20206	0.04882	4.14	0.000
Personal services	0.36007	0.07846	4.59	0.000
Social services	-0.65172	0.08530	-7.64	0.000
Local variables				
Mountain area (km2)	-0.00012	0.00008	-1.57	0.116
University centre	0.20579	0.06424	3.20	0.001
% Foreigner	-0.01301	0.01436	-0.91	0.365
Average firm size	0.01810	0.01007	1.80	0.072
Industrial district	-0.10793	0.05716	-1.89	0.059
(log) Employment density	-0.08445	0.03447	-2.45	0.014
% Upper Secondary and Tertiary degree	-0.00746	0.00634	-1.18	0.239
Accident/1000 cars	0.00959	0.00939	1.02	0.307
Local unemployment rate	-0.01680	0.01516	-1.11	0.268
Constant	1.32932	0.45201	2.94	0.003

Table 3.a – Probability of employment (selection equation)

Variables	Coef.	Robust Std. Err.	z	P> z
Individual variables				
Age	-0.04715	0.00919	-5.13	0.000
Age square	0.00059	0.00012	5.04	0.000
<i>Education (ref. secondary ed.)</i>				
Tertiary education	-0.23448	0.02852	-8.22	0.000
Post tertiary education	-0.51749	0.12232	-4.23	0.000
<i>Municipality size (ref. 0-10,000)</i>				
10,001 – 100,000	0.06026	0.05068	1.19	0.234
More than 100,000	-0.07316	0.17323	-0.42	0.673
<i>Gender/Marital status/location (ref. City man single)</i>				
City woman married	-0.20895	0.07094	-2.95	0.003
City man married	0.82936	0.05982	13.86	0.000
Rural woman married	-0.80194	0.22102	-3.63	0.000
Rural man married	0.60417	0.22788	2.65	0.008
City woman single	0.07855	0.04789	1.64	0.101
Rural woman single	-0.27376	0.21663	-1.26	0.206
Rural man single	-0.03888	0.21949	-0.18	0.859
Children 0-12	-0.30831	0.02813	-10.96	0.000
High marks	0.12458	0.02190	5.69	0.000
Regular educational path	0.17668	0.02432	7.26	0.000
Internet	0.43635	0.03263	13.37	0.000
English	0.02545	0.02307	1.10	0.270
<i>Father education (max lower secondary)</i>				
Upper secondary	-0.07635	0.02571	-2.97	0.003
Tertiary degree	-0.23728	0.04630	-5.12	0.000
Foreigner	-0.18094	0.07404	-2.44	0.015
<i>Macroarea (North West)</i>				
North East	-0.03108	0.21968	-0.14	0.888
Centre	-0.19995	0.23822	-0.84	0.401
South	-0.04856	0.29051	-0.17	0.867
Islands	0.03197	0.38209	0.08	0.933
Local variables				
Mountain area (km2)	0.00009	0.00020	0.47	0.637
University centre	0.10555	0.15122	0.70	0.485
% Foreigner	0.01960	0.04397	0.45	0.656
Average firm size	0.01970	0.01955	1.01	0.314
Industrial district	0.12809	0.14435	0.89	0.375
(log) Employment density	-0.00723	0.07990	-0.09	0.928
% Upper Secondary and Tertiary degree	-0.00362	0.02007	-0.18	0.857
Accident/1000 cars	-0.00346	0.02959	-0.12	0.907
Local unemployment rate	-0.08523	0.04041	-2.11	0.035
Selection variable				
Rent/Loan for accommodation	0.31679	0.02427	13.05	0.000
Constant	1.04521	0.63751	1.64	0.101
/athrho				
/athrho	-0.16393	0.29277	-0.56	0.576
rho				
rho	-0.16247	0.28505		
Wald test of independent equations (rho = 0): chi2(1) = 0.31 Prob > chi2 = 0.5755				

Table 3.b – Probability of overeducation (with selection bias)

Variables	Coef.	Robust Std. Err.	z	P> z
Individual variables				
Age	0.02720	0.02324	1.17	0.242
Age square	-0.00036	0.00030	-1.19	0.233
Specific experience	-0.01765	0.01076	-1.64	0.101
Specific experience square	0.00024	0.00029	0.83	0.406
Generic experience	-0.00171	0.01070	-0.16	0.873
Generic experience square	0.00011	0.00027	0.39	0.693
<i>Education (ref. secondary ed.)</i>				
Tertiary education	0.01396	0.09887	0.14	0.888
Post tertiary education	-0.32637	0.26909	-1.21	0.225
Permanent contract	-0.10398	0.05169	-2.01	0.044
Full time	-0.22430	0.05676	-3.95	0.000
<i>Municipality size (ref. 0-10,000)</i>				
10,001 – 100,000	-0.02907	0.05367	-0.54	0.588
More than 100,000	-0.00790	0.07988	-0.10	0.921
<i>Gender/Marital status/location (ref. City man single)</i>				
City woman married	-0.11477	0.10286	-1.12	0.265
City man married	-0.15605	0.12785	-1.22	0.222
Rural woman married	-0.05477	0.18493	-0.30	0.767
Rural man married	-0.20731	0.14349	-1.44	0.149
City woman single	-0.01383	0.13982	-0.10	0.921
Rural woman single	-0.03532	0.12733	-0.28	0.781
Rural man single	-0.02366	0.11187	-0.21	0.833
Children 0-12	-0.02704	0.08050	-0.34	0.737
High marks	-0.24830	0.04487	-5.53	0.000
Regular educational path	-0.10581	0.04567	-2.32	0.021
Internet	-0.45267	0.09514	-4.76	0.000
English	-0.16920	0.04650	-3.64	0.000
Daily commuting time (home to work)	-0.00287	0.00098	-2.92	0.003
Movers	-0.20203	0.07876	-2.57	0.010
<i>Father education (max lower secondary)</i>				
Upper secondary	-0.03883	0.05493	-0.71	0.480
Tertiary degree	0.06066	0.09300	0.65	0.514
Foreigner	0.78731	0.14492	5.43	0.000
<i>Macroarea (North West)</i>				
North East	-0.08536	0.05745	-1.49	0.137
Centre	0.15185	0.06376	2.38	0.017
South	0.11399	0.10345	1.10	0.270
Islands	-0.00514	0.13759	-0.04	0.970
Firm characteristics				
<i>Firm size (small: ma10-49)</i>				
Medium (50-499)	0.13846	0.05163	2.68	0.007
Large (500 or more)	0.09712	0.07756	1.25	0.211
<i>Branch of economic activity (ref. production of goods)</i>				
Production services	-0.10455	0.05753	-1.82	0.069
Distribution services	0.20009	0.04961	4.03	0.000
Personal services	0.35590	0.07895	4.51	0.000
Social services	-0.64582	0.08801	-7.34	0.000
Local variables				
Mountain area (km2)	-0.00013	0.00008	-1.54	0.124
University centre	0.19300	0.07333	2.63	0.008
% Foreigner	-0.01396	0.01482	-0.94	0.346
Average firm size	0.01601	0.01106	1.45	0.148
Industrial district	-0.12001	0.05963	-2.01	0.044
(log) Employment density	-0.08364	0.03370	-2.48	0.013
% Upper Secondary and Tertiary degree	-0.00728	0.00703	-1.03	0.301
Accident/1000 cars	0.00976	0.00980	1.00	0.319
Local unemployment rate	-0.00741	0.02301	-0.32	0.747
Constant	1.35395	0.45325	2.99	0.003

Annex 2 - Data and variables description

Dependent variable:

Overeducation, Isfol Plus 2005, dummy variable built on question v430_1: Is your educational degree necessary to perform your job?"; no = 1, yes = 0

Individual and firm-specific variables:

Age: Isfol Plus 2005, question V70_1

Specific experience: Isfol Plus 2005, calculated as: 2005 - starting year of current job (question V460_1)

Generic Experience: Isfol Plus 2005, calculated as: 2005 - starting year of first job (question v473_2)

Individual level of education: Isfol Plus 2005, questions V880_1 – V885_6

Permanent contract: Isfol Plus 2005, question V110_2

Full time: Isfol Plus 2005, question V140

Municipality size: Isfol Plus 2005, question V9280

Gender/Marital status/location: Isfol Plus 2005, question V80 (gender), question V970 (marital status), question V9142 (municipality type)

Children0-12: Isfol Plus 2005, dummy variable (children = 1; no children = 0) built on question V81

High marks: Isfol Plus 2005, dummy variable built on question V890 (lower secondary education; medium-high = 1, medium-low = 0); on question V892 (upper secondary education; from 60/60 to 48/60, or from 100/100 to 80/100, = 1, from 47/60 to 36/60, or from 79/100 to 60/100, = 0); and on question V898 (tertiary education; from 100 cum laude to 99 = 1, from 98 to 66 = 0).

Regular educational path: Isfol Plus 2005, dummy variable built on question V910 (tertiary education; tertiary education degree "in corso" or "fuori corso" within 3 years = 1; "fuori corso" longer than 3 years = 0); and on question V920 (lower and upper secondary education; no failure = 1; one, two or more failures = 0)

Internet: Isfol Plus 2005, dummy variable built on question V1090_2 (Can you search information on the Web?)

English: Isfol Plus 2005, dummy variable built on question V1090_4 (Can you make a telephone conversation in English?)

Daily commuting time (home to work): Isfol Plus 2005, variable V600_1 (How long do you take to get from home to work?)

Movers: Isfol Plus 2005, variable V430_2 (With reference to your current job, did you have to move for working reasons?)

Father education: Isfol Plus 2005, dummy variable built on question *titstupa*

Foreigner: Isfol Plus 2005, dummy variable built on question V1100 (Italians = 0; Foreigners = 1)

Firm size: Isfol Plus 2005, question V400_1; (small firms, from 0 to 49 employees; medium firms, from 50 to 499 employees; big firms, 500 or more employees).

Branch of economic activity: Isfol Plus 2005, dummies variables built on question *sett5*

Territorial dummies: Isfol Plus 2005, dummy variable built on question *area5*: North West (ref. category), North East, Centre, South, Islands.

Local variables:

Mountain area (km²): Atlante dei comuni, ISTAT, 2005.

University centre: Atlante dei comuni, ISTAT, 2005.

% of foreigners: Atlante dei comuni, ISTAT, 2005.

Average firm size: average number of employees in local units (elaboration on data from *Censimento dell'Industria* 2001, Istat).

Industrial district: dummy variable equal to 1 if the LLM corresponds to an industrial district, 0 otherwise, according to ISTAT classification (*Censimento dell'Industria* 2001, Istat).

(Log) Employment density: Log of employment per squared km (elaboration on data from *Censimento della Popolazione* 2001 and *Censimento dell'Industria* 2001, Istat).

% of upper secondary and tertiary graduates: resident population older than 6 years by level of education (*Censimento della Popolazione* 2001, Istat).

Accidents per 1000 cars: Atlante dei comuni, ISTAT, 2005.

Unemployment rate: *Censimento della Popolazione* 2001, Istat.

Selection variable:

Rent or loan for accommodation: Isfol Plus 2005, dummy variable built on questions V1671 and V1670_1.

References

- Baum S., Bill A., Mitchell W., 2008, Labour Underutilisation in Metropolitan Labour Markets in Australia: Individual Characteristics, Personal Circumstances and Local Labour Markets, *Urban Studies*, 45(5, 6) May, 1193–1216.
- Bratti M., Leombruni R., 2009, Local human capital externalities and wages at the firm level: the case of Italian manufacturing, IZA Discussion Paper, 4613.
- Brunello G., De Paola M., 2008. Training and economic density: some evidence from Italian provinces. *Labour Economics* 15(1): 118-140.
- Büchel F., Battu H., 2003, The Theory of Differential Overqualification: Does It Works?, *Scottish Journal of Political Economy*, Vol. 50, No 1, February, 1-16.
- Büchel F., Van Ham M., 2003, Overeducation, Regional Labour Markets, and Spatial Flexibility, *Journal of Urban Economics* 53, 482-493.
- Cahuzac E., di Paola V., 2004, Overeducation and Wages Downgrading: Conditions of Spatial Differentiation, Working Paper presented to the 16th SASE Annual Meeting on Socio-Economics, George Washington University, Washington D.C. USA, July 8-11, 2004.
- Croce G., Ghignoni E., 2009, *Employer-provided training and knowledge spillovers Evidence from Italian local labour markets*, Working Papers, Dipartimento di Economia Pubblica, University of Rome La Sapienza, n. 130, November.
- Dalmazzo A., De Blasio G. (2007), Social returns to education in Italian local labour markets, *Annals of regional sciences*, 41, 51-69.
- Dell'Aringa C., Pagani L., 2010, Labour Market Assimilation and Over Education: The Case of Immigrant Workers in Italy, Working Paper DISCE No 58, Università Cattolica del Sacro Cuore, Milan, January.
- Di Addario S., Patacchini E. (2008), Wages and the city. Evidence from Italy, *Labour Economics*, 15, 1040-1061.
- Duncan G., Hoffman S., 1981, "The Incidence and Wage Effects of Overeducation" *Economics of Education Review*, 1(1), pp.75-86.
- Frank R.H., 1978, Why women earn less: The theory and estimation of differential overqualification, *American Economic Review* 68, 360–373.
- Franzini M., Raitano M., 2009, "Pochi e poco utilizzati? L'overeducation dei laureati italiani", Working Paper presented to the 50th SIE Scientific Annual Meeting, LUISS, Rome, October 22-24, 2009.
- Ghignoni E., 2001, "Frontiere di competenza, overeducation e rendimento economico dell'istruzione nel mercato del lavoro italiano degli anni '90", *Rivista di Politica Economica*, F. VII-VIII, pp 115-158, July-August.
- Hartog J. (2000), Over-education and earnings: where are we, where should se go?, *Economics of education review*, 19, pp. 131-147.
- Hartog J., Oosterbeek H., 1988, "Education, Allocation and Earnings in the Netherlands: Overschooling?", *Economics of Education Review*, 7(2), pp. 185-194.
- Hensen M., de Vries R., 2004, The relationship between geographical mobility and education-job mismatches, ERSA conference papers ersa04p158, European Regional Science Association.
- Jensen U., 2003, "Measuring Overeducation with Earning Frontiers and Panel Data", *Overeducation in Europe. Current Issues in Theory and Policy*, Büchel F., de Grip A., Mertens A. (eds.), Edward Elgar, Cheltenham, UK; Northampton, MA, USA.
- Kiker B., Santos M., Mendes de Oliveira M., 1997, "Overeducation and Undereducation: Evidence for Portugal", *Economics of Education Review*, 16(2), pp. 111-125.
- Linsley I., 2005a, Overeducation in the Australian Labour Market: Its Incidence and Effects, Research Paper, University of Melbourne, Department of Economics, Melbourne Victoria, Australia, No 939, June.
- Linsley I., 2005b, Causes of Overeducation in the Australian Labour Market, Research Paper, University of Melbourne, Department of Economics, Melbourne Victoria, Australia, No 940, June.
- McGuinness S., (2006), Overeducation in the labour market, *Journal of economic surveys*, 20, 3, pp. 387-418.
- Moulton, B.R., 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, Vol. 72(2): 334-338.
- Ordine, P., Rose, G., 2009a, "Overeducation and Instructional Quality: a Theoretical Model and Some Facts", *Journal of Human Capital*, Vol. 3, No 1, 73-105.

- Ordine, P., Rose, G., 2009b, "Inefficient self-selection into education and wage inequality in Italy", Working Paper presented to the XXIV National Conference of Labour Economics, University of Sassari, 24-25 September 2009
- Preston V., McLafferty S., 1999, Spatial mismatch research in the 1990s: Progress and potential, *Papers in Regional Science* 78, 387-402.
- Rumberger R., 1987, "The Impact of Surplus Schooling on Productivity and Earnings", *Journal of Human Resources*, 22(1), pp. 24-50.
- Sanroma E., Ramos R., 2004, Overeducation and Local Labour Markets in Spain, Working Paper presented to the 44th European Regional Science Association Congress, Porto, 25-29 August 2004.
- Sattinger M., 1993, Assignment models of the distribution of earnings, *Journal of economic literature*, 31, pp.831-880.
- Sloane P.J. (2003), Much ado about nothing? What does the overeducation literature really tell us?, in Buchel F., de Grip A., Mertens A. (eds), *Overeducation in Europe. Current issues in theory and policy*, Edward Elgar, Cheltenham, UK, Northampton MA, USA, pp. 11-48.
- Thurow L.C., 1975, *Generating inequality*, Basic Books, New York.
- Van de Ven W., Van Praag B., 1981, "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection", *Journal of Econometrics*, n. 17, pp. 229-252.
- Van Ham M., 2002, *Job Access, Workplace Mobility, and Occupational Achievement*, Eburon, Delft.
- Verdugo R., Verdugo N., 1988, "The Impact of Surplus Schooling on Earnings", *Journal of Human Resources*, 22(4), pp. 690-695.