View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Munich Personal RePEc Archive

Overeducation and spatial flexibility in Italian local labour markets

Croce, Giuseppe and Ghignoni, Emanuela
Department of economics and law, Sapienza University of Rome

01. March 2011

Online at http://mpra.ub.uni-muenchen.de/29670/
MPRA Paper No. 29670, posted 17. March 2011 / 17:01

Overeducation and spatial flexibility in Italian local labour markets

Giuseppe Croce[a], Emanuela Ghignoni[b]

[a]Department of Economics and Law, Sapienza University of Rome,
Via del Castro Laurenziano 9, 00161 Rome, Italy; e-mail: giuseppe.croce@uniroma1.it
[b]Department of Economics and Law, Sapienza University of Rome,
Via del Castro Laurenziano 9, 00161 Rome, Italy; e-mail: emanuela.ghignoni@uniroma1.it

Abstract

According to a recent strand of literature this paper highlights the relevance of spatial mobility as an explanatory factor of the individual risk of being overeducated. To investigate the causal link between spatial mobility and overeducation we use individual information about daily home-to-work commuting time and choices to relocate in a different local area to get a job. In our model we also take into account relevant local labour markets features. We use a probit bivariate model to control for selective access to employment, and test the possibility of endogeneity of the decision to migrate. Separate estimations are run for upper-secondary and tertiary graduates.

The results sustain the appropriateness of the estimation technique and show a significantly negative impact of the daily commuting time for the former group, as well as, negative impact of the decision to migrate and of the migration distance for the latter one.

Keywords: Overeducation, Spatial flexibility, Local labour markets, Sample selection bias

JEL classification: J21, J61, J62

1. Introduction

Labour market theories often assume that workers search for job in the global labour market, whilst most of them actually stay within the local one. If this geographical constraint is taken into account, the spatial mismatch hypothesis can explain the presence of unemployment with the lack of local job opportunities in combination with limited spatial flexibility of the workforce (Preston & 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, including overeducated workers (Hensen & de Vries, 2004; Baum, Bill & Mitchell, 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 influence of spatial effects on underemployment, and in particular on overeducation, have been neglected for long, with the commendable exception of the theory of differential overeducation concerning the job seeking behaviour of married women (Frank, 1978; Büchel & Battu, 2003). In this framework, the seminal paper of Büchel & van Ham (2003) highlighted the role of workers' spatial mobility and local labour markets (LLMs) characteristics as potential explanatory factors of overall overeducation and not only of the differential overeducation of married women.

Following this strand of literature, we estimate a model where worker's spatial mobility is assumed to affect the individual risk of overeducation. The contribution of this paper is twofold. Firstly, compared to previous papers, it adds a few innovations to the model specification. Indeed we look into the effects of commuting time and decision of moving for work reasons as a measure of spatial flexibility. Secondly, to the best of our knowledge, this is the first attempt to evaluate the influence of spatial mobility on overeducation based on Italian data. Thus, our results broaden the range of hypotheses that have been empirically tested so far to explain overeducation in Italy. Moreover, our study is based on a representative sample of the stock of both university as well as high school graduates whereas previous papers mostly consider recent cohorts of university graduates. This allows us to bring new evidence on the set of factors shaping the risk of overeducation. Among these factors we include a number of variables besides the local unemployment rate, which is usually considered, to characterise the LLM where the worker resides. This way, we intend to ascertain the possible effect of the LLMs characteristics, as the territorial dimension has proved to be an important factor in the literature analysing a wide range of individual outcomes in the Italian labour market (Cannari, Nucci & Sestito, 2000; Dalmazzo & De Blasio, 2007; Di Addario & Patacchini, 2008).

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

2. Theoretical framework

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 the aggregate amount of workers with a tertiary degree or a high school diploma exceeds the amount of available jobs requiring such levels of education. If wages are not flexible enough or firms do not adapt their jobs to workers' characteristics, such a mismatch will tend to persist.

Possible explanations in this line are given by the job competition and job assignment theories (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 taken into account 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 affect the distribution of the risk of being overeducated among heterogeneous workers. Following Frank (1978), McGoldrick and Robst (1996), and Büchel and van Ham (2002), the spatial flexibility hypothesis focuses, in particular, on explanations based on 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 effect of the amount of job opportunities in LLMs (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), in Büchel and Van Ham (2003) geographical restrictions play a key role in explaining overeducation, but their effect concerns all workers (and not only women) and the possibility of extending the job search to other LLMs through commuting is also considered. Their results for the German labour market show that variables affecting the individual spatial flexibility or related with the spatial distribution of employment can explain overeducation. In particular, a higher individual mobility (owning a car) permits to enlarge the "effective" size of the labour market, which decreases the probability of being overeducated, although this result, as the authors recognise might be due to endogeneity as the overeducated are less likely to be able to afford a motor vehicle if overeducation is negatively correlated with earnings. By contrast, the local unemployment rate, which can be regarded as an index of competition for jobs in the LLM, has a significant (negative) impact on the probability of being employed but it does not affect the probability of being overeducated.

Along this line, Sanroma and Ramos (2004) analyse the possible effects of a number of variables related to LLMs in Spain and take into account the possibility of widening the spatial job search through the inclusion in the model of two variables capturing the availability of private transport (at individual level) and the number of road kilometres per car (at provincial level). Their results, in line with those of Büchel and van Ham (2003), show that living in a small town increases the probability of being overeducated, whilst the possibility of searching for a job in wider labour markets significantly reduces it.

In a study on French data (Cahuzac & di Paola, 2004) the authors include in the model the average travelling time from home to work at local level, which should account for geography and the quality of local transport infrastructures. Moreover, they use a series of dummies based on the notion of "poles urbains" (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, in denser labour markets a large availability of job offers can increase the probability of finding a well-matched job; on the other hand, fiercer competition among workers in that areas could decrease the chance of finding a suitable position. Results show that the risk of overeducation increases with the average travelling time from home to work. Moreover, overeducation is lower in denser and more populated areas and in LLMs characterised by lower unemployment rates. Thus, an easier access to a large concentration of jobs seems to lead to better matches.

In an analysis of overeducation in the Australian labour markets, Linsley (2005a) assumes that spatial constraints are more binding for married women with children and individuals without access to adequate transport options or residing in locations more distant from urban labour markets. In practise, the access to a motor vehicle is used to capture the individual's capacity to commute, while some location variables (major city, inner regional areas, outer regional and remote areas) are included to capture the size of LLM and the distance between individual's residence and

agglomerations of jobs. The analysis concludes that women (whether married or not) living in outer regional or remote areas face a significantly higher risk 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.

As for Italy, about 30% of university graduates declare that their degree is not required to perform their job, three years after completing their studies , and this has decreased slightly from 2001 to 2007[1]. Even though this is a sizeable figure, it cannot be concluded that it signals an excess of highly educated labour supply in Italy as, despite its recent increase, the proportion of university graduates in Italy remains substantially lower than in comparable countries. Moreover, the return on education investment is not higher than in other developed countries (Brunello, Comi & Lucifora, 2001). This evidence points out a lack of relative demand for highly educated workers and a widespread qualitative mismatch between workers and jobs.

Among studies on overeducation in Italy, Di Pietro and Urwin (2006) in a paper based on a sample of individuals graduated in 1998 and interviewed after three years, find that those overeducated earn less than their peers whose educational level fits their job profile. They argue that, faced with the recent increase in the supply of graduates in the Italian labour market, employers could have taken advantage of this by raising the educational level required to get a job without a corresponding improvement of the wage paid for that job. The evidence of the wage penalty due to overeducation is confirmed also by Franzini and Raitano (2011) who exploit a sample drawn from a different survey.

Di Pietro and Cutillo (2006) test the hypothesis of a causal relationship between the quality of the attended university and the risk of being overeducated. Their results confirm that graduates from high quality research institutions face a lower risk of being overeducated. In a similar vein Ordine and Rose (2009a,b), by adopting different measures of quality, find that graduating from research oriented universities significantly reduces overeducation.

Overall, these papers focus on individual and job features as determinants of the likelihood of overeducation but fail to take into account the impact of individual labour mobility. This paper aims at filling this gap.

3. Methodology and data

3.1. Measuring overeducation

A worker can be considered as overeducated when he/she possesses a level of education in excess of that which is required for his/her particular job. Empirical literature suggests various possible measures of overeducation, among which the most commonly used are the *objective* measure (Rumberger, 1987), the *subjective* measure (Duncan & Hoffman, 1981; Hartog & Oosterbeek, 1988) and the *empirical* measure (Verdugo & Verdugo, 1988; Kiker, Santos, & Mendes de Oliveira, 1997). Unfortunately, these methods provide rather different results when applied to the same database, and the application of one or the other measurement is generally determined by the available information.

In this framework, Isfol-Plus survey can be regarded as a useful source of information in order to analyse overeducation in Italy (Franzini & Raitano, 2011). We measure overeducation by means of the (direct) subjective method, as the overeducated are identified through the question *Is your educational degree necessary to perform your job?* Consequently, a few caveats that are common to all overeducation analyses based on subjective measures, have to be kept in mind. Obviously, respondents do not specify whether they are referring to a formal or substantial necessity and they do not indicate the to what extent they actually use the skills acquired throughout their educational

[1] This share amounted to 32,5% in 2001 and dropped to 32.1% in 2004 and to 28% in 2007 (this last figure is referred to those who completed the three-years courses) (Istat, 2003, 2006 and 2010).

path. Moreover, it is difficult to distinguish genuine cases of overeducation (jobs below the individual educational level) from mismatches due to having a job outside the individual field of education. However, the large size of the sample and the rich set of information provided by the survey allow controlling for a variety of original explanatory variables, including a few proxies for individual abilities (school grades, regular schooling paths) and practical skills (in particular, computer and foreign language skills).

3.2. Data

To carry out our analysis we exploit individual data drawn from the 2005 wave of the Isfol-Plus survey and data from various sources on LLMs[2]. We select individuals with an upper secondary or a tertiary degree (26,323 individuals), aged 20 years or more (25,522 individuals), who declare to be employed or not employed, with the exception of students and retired people, which leads to a further sample reduction (19,335 individuals, 12,694 of which are employed). We include in the analysis both dependent and independent employees, as well as upper secondary and tertiary graduates. Moreover, our study is based on a stock of population (as in Büchel & Van Ham, 2003) and not on the flow of new graduates (as in part of the most recent literature). Indeed, this allows us to gain a broader view of the determinants of overeducation.

We first estimate a pooling baseline model where a dummy variable distinguishes upper secondary, tertiary and post-tertiary graduates in order to ascertain whether they face a different risk of overeducation (see Table 3, col. 1). Then, we proceed by performing separate estimations for the upper secondary, on one hand, and for the tertiary/post-tertiary graduates on the other hand, as these two groups may differ from each other in the constant terms and in the coefficients (cols. 2 and 3).

3.3. The sample selection bias problem

The objective of the paper is to identify factors explaining the probability of being overeducated, keeping into account individual and job specific characteristics, as well as a few variables related to LLMs. Among individual factors we look into the effects of individual spatial flexibility.

To achieve valid results our empirical analysis had to cope with two main methodological issues concerning sample selection and endogeneity of workers' change of residence. First of all, the estimations may suffer from a sample selection problem, as overeducation is observable only if the individual actually works (Ordine & Rose, 2009a; Büchel & Van Ham, 2003). As well-known, analysing overeducation while restricting the sample to the employed could lead to biased results. Indeed, there could be some unexplained factors that positively affect the probability of being in employment and at the same time are negatively correlated with the risk of overeducation (Di Pietro & Cutillo, 2006). In this case those least likely to enter employment would be the most likely to be overeducated, and inactivity could be chosen as a strategy to avoid overeducation. By contrast, some omitted factors positively affecting the probability of being employed, such as aversion to unemployment risks, could be positively related to the probability of being overeducated. In this case, some people would settle for jobs for which they are overeducated in order to avoid unemployment. To deal with this problem we estimate the following bivariate probit model with sample selection (Van de Ven & Van Praag, 1981):

$$P_{OVER} = \alpha + \beta X + \delta Y + \varepsilon_1$$
$$P_{WORK} = bZ + \varepsilon_2 \qquad [1]$$

[2] For the purpose of this paper LLMs 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).

where

$P_{OVER} = \begin{cases} 1 & \text{if the individual is overeducated} \\ 0 & \text{otherwise} \end{cases}$ and $P_{WORK} = \begin{cases} 1 & \text{if the individual actually works} \\ 0 & \text{otherwise} \end{cases}$

X and *Y* are, respectively, a vector of individual and job-specific characteristics, and a vector of LLM characteristics influencing the probability of overeducation, whilst *Z* is a set of variables that affect the probability of working.

In particular, amongst individual characteristics we include some proxies for individual unobservable abilities (measures of school performances, such as grades and regularity of educational paths) and some transversal skills (surfing the Internet and speaking foreign languages). Unfortunately, we do not have information about the field of study for tertiary graduates. By contrast we dispose of information about the type of upper secondary degree, thus we can include a dummy pointing out upper secondary graduates with technical versus academic (*liceo*) degree. Job specific characteristics include the branch of economic activity, firm size and type of contract.

As for individual spatial flexibility we use two mobility indicators: daily commuting time (*How many minutes do you take to get from home to work?*) and a dummy for movers (*With reference to your current job, did you have to move for working reasons?*) (see Appendix 2 for details). To take into account some possible interactions between workers' mobility and family ties, as well as some specific features of the Italian regional labour markets, we interacted both the mobility indicators with gender and the "movers" dummy with the "direction" of the residence change for working reasons (from a southern Province to a northern one). Nevertheless results have proved to be not significant and they are not showed here in order to save space. Finally, to evaluate the influence of the distance covered by individuals who moved for working reasons, we substituted the movers dummy with the road distance (in kilometres and kilometres squared) between the Province of origin and the present Province of residence.

To complete the picture, we take into account some LLM characteristics which may affect overeducation risks. In particular, we include in the model some proxies for the easiness of communication and for the quality of transportation infrastructures at local level (the extension in squared kilometres of mountainous areas in the LLM and the number of accidents per 1,000 cars in circulation); some indicators of the local human capital (a dummy indicating the university towns and the percentage of upper secondary and tertiary graduates in the local population); and some indicators of the production structure at local level (the average firm size and a dummy for LLMs corresponding to industrial districts). In this framework, the percentage of foreign immigrants at local level has been included in order to control for the possibility that some low-skill jobs mostly rely upon immigrant workers. Lastly, we control for some indicators of competition among workers. In particular, we speculate that a high level of unemployment could worsen the individual chance of finding a suitable job. Moreover, we control for local employment density (the number of employed per squared kilometre) that, as argued above, can have two opposite effects on overeducation risk. On the one hand, the large number of job offers in denser LLMs could increase the probability of finding a suitable job. On the other hand, employment density could lower the chances of finding a job matching workers' skills due to a stronger competition among them. Thus, the sign of the correlation between employment density and the probability of overeducation is not known *a priori* and it depends on which effect would prevail.

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

The identification of the bivariate probit model with sample selection, reported in [1], is based on the choice of at least one selection variable that affects only the access to employment, but not the

overeducation risk. We exploit the information on individuals who have to pay a mortgage loan for accommodation to build our selection variable. Actually, we expect that this financial burden is strongly (positively) correlated with the probability of being employed; whereas, conditional on working, it should not have a direct influence on labour market outcomes once we control for individual and job-specific characteristics, and local variables. The question about the loan in the Isfol-Plus survey (*Do you pay a loan for your accommodation?*) detects whether a loan weighs upon the respondent's household. In this regard, we assume that the likelihood of being employed increases with the incidence of individual income on the household income. Thus, we build a variable by weighting a dummy which distinguishes between those with a loan and those without it, by the share of the respondent's individual income on the total income of his/her own household (see Appendix 2 for details). As it will be shown below, we adopt this weighted loan variable in the estimates concerning upper secondary graduates, while we use the simple loan dummy in the case of university graduates, due to the large number of missing values concerning the individual and household incomes.

3.4. The endogeneity of migration

A further econometric problem could arise from the possibility of endogeneity of the individual's choice to relocate to get a job (Dostie & Leger, 2009). More precisely, there could be a potential correlation between overeducation and unobserved individual characteristics influencing the decision to migrate for working reasons. We deal with this problem in two ways.

Firstly, as mentioned above, some proxies for unobservable individual abilities (such as having reported high school/university marks and having concluded the educational path in time) are included in the overeducation equation. Secondly, we implement the Hausman t-test (Hausman, 1978; Davidson & MacKinnon, 1993) to check for endogeneity.

To this aim we have to select a valid instrument, i.e. a variable that significantly influences the migration choice but can be legitimately excluded from the overeducation equation, conditional on the other explanatory variables. Following the relevant literature (Audas & Dolton, 1999), we chose a measure of the labour market conditions in the area where the mover comes from. In particular, we used the long-term unemployment rate of the province of origin (average value from 1981, 1991 and 2001 Censuses data) assuming that people living in provinces with a worse structural employment situation are more likely to migrate with respect to those living in provinces with a better situation. On the other hand it seems reasonable that the lagged unemployment rate of the province of origin has no influence on the present risk of overeducation.

Then, firstly a probit migration equation where the instrument has been included is estimated and, secondly, the residuals from the first stage regression are added as an explanatory variable in the overeducation equation. If the t-test accepts the null hypothesis that the estimated coefficient on the residuals is not significantly different from zero, one may conclude that the suspected variable (movers) is not endogenous and that a regression technique without instrumental variables is appropriate to estimate the overeducation function. This condition was satisfied for both upper secondary and tertiary graduates (the unemployment rate of the province of origin shows a strong positive influence on the migration probability, and the Hausman test produces a t-statistics of -0.72 for the pooling sample, -1.41 for upper secondary graduates and -0.32 for tertiary graduates).

4. Estimations results

Some descriptive characteristics are provided in Table 1, which reports information not only on the full sample but also on the employed alone, which represent two thirds of the full sample. In particular, it can be observed that, on average, the employed spend about 21 minutes commuting every day, while 12% of them have moved for working reasons.

From Table 2, which depicts the incidence of overeducation in the sample, we can see that upper secondary graduates declare to be overeducated more often than tertiary graduates. This fact sounds rather surprising as the risk of overeducation is expected to increase with the level of education. In addition, workers with an upper secondary degree display on average a higher mobility as 35% of them spend more than 30 minutes commuting and one third of them have moved for working reasons, compared to 18% and 13% of the university graduates respectively.

TABLE 1 AND TABLE 2 ABOUT HERE

Firstly, we estimated a baseline model where the probability of being overeducated depends on a set of determinants plus our key-variables indicating commuting time and movers. In the rest of the paper we will run separate estimations for upper secondary and university graduates in order to gain insights into the differential impact of spatial flexibility on their outcomes in the labour market. Indeed they can be regarded as quite different groups, and besides the two samples are large enough to allow us to run distinct regressions. Nevertheless, we also display in the first column of Table 3 the results of a regression on the overall sample, encompassing both high school and university graduates, as this estimation allows ascertaining whether these two groups face a different risk of overeducation. Our findings confirm that, even after controlling for all other characteristics, those holding a university or a post-university degree are less likely to be overeducated than high school graduates (the same result was found by Chevalier, 2003, on UK data). Moreover, both variables measuring spatial mobility have a negative and significant coefficient, which suggests that mobility is associated to a lower risk of overeducation.

Nevertheless, when the same model is estimated on the two subsamples, two distinct patterns arise. Indeed, as for high school graduates, we find that a longer commuting time is associated with a lower risk of overeducation, while having moved for working reasons does not affect the quality of their current match (Table 3, col. 2). On the other hand, turning to the university graduates (Table 3, col. 3), it emerges that movers are significantly less overeducated whereas the time spent in commuting does not alleviate overeducation (the coefficient is still negative but it is not significant).

TABLE 3 ABOUT HERE

However, these results cannot be considered robust until the problem of sample selection bias has been addressed. To this aim, the next step was the estimation of the bivariate probit model. Actually, as for the sample of the high school graduates, we can see from Table 4.a that the results are substantially confirmed even though the statistical significance of a few variables decreases compared to the baseline model. On the other hand, the results concerning the university graduates are almost identical to the previous ones. Moreover, what is most important is that in either case the sign and the statistical significance of our key-variables are validated. In particular, even after controlling for selection bias, it is confirmed that the longer the time spent in commuting the lower the risk of being overeducated. This evidence confirms our hypothesis of spatial flexibility predicting that workers able to search for job in a larger area are more likely to benefit from their own human capital to the fullest. In accordance with this hypothesis, individuals who accept longer daily commuting times bear lower overeducation risks. Thus, individual mobility is recompensed with a better skills-job match. More in general, this result points out that frictions and barriers in the labour market can offer a possible explanation of overeducation apart from imbalances between the demand and supply of qualified labour at an aggregate level.

TABLES 4a AND 4.b ABOUT HERE

Among the other results it can be noted that the variable age is not fully significant (albeit positive), and its effect is not linear but decreasing as shown by the negative and significant coefficient of the age squared variable. Instead, the effect of tenure, which was negative and significant in the baseline model, loses its statistical significance in the bivariate probit[3]. The presence of children aged 0-12 and the size of population in the area are not relevant.

Holding a temporary contract shows a positive coefficient as in the baseline model, however it is no longer statistically significant. Thus, it can be concluded that the duration of the employment contract does not alter the probability of being overeducated. Overall, we cannot say that the workers in a fixed term job trade off employment stability for a lower risk of overeducation. On the contrary, this evidence can be read as a signal of dualism in the labour market as it points out that workers in temporary employment suffer a worse match too.

As for the marks reported at the end of school and the regularity of the school path (having finished high school in the regular time), which can be regarded as proxies of unobserved ability, our results show that both reduce overeducation (but the second one is only weakly significant). Even the possession of general skills such as the ability of surfing the Internet and speaking English result to be helpful to lessen the risk of being overeducated. Among people with a secondary education level, those who got their degree in a technical school are less overeducated than those who graduated in a more academic high school (*licei*). Indeed, the negative coefficient of the dummy "technical skills" seems to point out that such skills are effective in fostering individual employability. However the dummy could also capture some unobserved adverse characteristics of those who come from a *liceo* and give up a university degree or failed to get it. Foreigners occupy more often jobs for which their educational degree is not necessary[4], whereas working in the public sector appears to imply a higher probability of an adequate match.

We added to the model a series of interactions between gender, marital status and the type of area of residence (rural/urban) to control for the presence of differential overeducation (Frank, 1978; Büchel & Battu, 2003). Actually, it could be argued that married women in rural areas should suffer higher risk of overeducation. Nevertheless, as in Büchel and Battu (2003), none of these interaction terms is significant.

Finally, our results show that a high school graduate living in a LLM corresponding to an industrial district faces a lower risk of overeducation. This suggests that in the industrial districts, which are characterized by a stable core of economic activities, the educational structure of labour demand and supply tend to be more homogeneous. By contrast, as found by Büchel and Van Ham (2003), we do not find any evidence of a dependence of overeducation on the local unemployment rate[5].

Looking at the employment equation (Table 4.b), it is worth noting that, in line with our hypothesis, the selection variable affects the probability of being employed. The effect of the loan dummy is positive and highly significant: people paying a loan are more likely to be in employment. At the same time we verified by a separate probit equation (not shown here to save space) that the same variable does not affect overeducation. Then it matches the requirements in order to be a proper selection variable. Moreover, the *rho* coefficient is negative but not statistically significant, meaning that our sample is not severely affected by a sample selection bias.

The probit bivariate model was estimated for the tertiary graduates too. Even in this case it is confirmed that spatial flexibility matters. In particular, it turns out that movers face a lower risk of overeducation (Table 5.a). It is worth noticing that this effect continues to be as strong and statistically significant as it was in the baseline model. On the other hand, time spent on commuting is not helpful in countering overeducation.

---

[3] Since age and tenure may be correlated, we also estimated our models by excluding, in turn, age and specific experience. Results are stable.

[4] Note that in German labour markets, analysed by Büchel and Van Ham (2003), being foreigner do not significantly affects the probability of being overeducated.

[5] See, contra, Cahuzac and di Paola (2004) on French local labour markets.

The coefficients of age as well as tenure are far from being statistically significant[6]. Then, we don't find any evidence that overeducation tends to disappear with age and experience. This is in line with Büchel and Van Ham (2003) who find that age does not influence overeducation. Although longitudinal data would be required in order to detect the effects of time, such a result suggests that overeducation can represent a persistent condition. This contradicts career mobility theories (Linsley, 2005a, 2005b), suggesting that workers who stay longer in their job are those in a more qualified position, facing a lower probability of overeducation. This also implies that accumulation of specific human capital through experience and training within a firm does not play a relevant role in the allocation of workers to jobs.

On the contrary, when university graduates hold a fixed term contract or are self-employed they are less likely to be overeducated than their peers in a permanent job. This evidence suggests that workers with a university degree are better able than high school graduates to trade-off employment stability for a proper matching according to their individual preferences.

Among the other individual characteristics, it can be noted that having graduated by the scheduled time is strongly associated with a lower probability of overeducation. This effect can be attributed to unobserved individual abilities influencing this probability. Equally, overeducation is less frequent for people reporting high marks at the end of university. Nonetheless, this effect is statistically weaker. On the contrary, skills in surfing the Internet and foreign languages are not helpful. Turning to family background measured on the basis of the educational level of the father, this results to affect the quality of the individual match. In particular, workers with a father holding an upper secondary or a tertiary degree are more likely to be in a proper job. Foreigners tend to be more overeducated than natives.

Among the variables measuring LLM features, the unemployment rate is far from being relevant. On the other hand, the number of accidents per thousands of cars, which can be seen as a measure of local congestion and of the quality of the transport infrastructures in the LLM, has a positive and strongly significant effect. Thus, we can regard this finding as a further indirect support to the hypothesis of spatial flexibility: workers living in areas where the costs of mobility are higher suffer a higher risk of overeducation. On the other hand, the share of people holding a tertiary degree in the LLM is associated with a lower risk of overeducation. This suggests that a higher density of university graduates signals better job opportunities for them rather than scarcer as it would be if the effect of competition among workers would prevail.

The employment equation shows that having to repay a home loan does affect the probability of employment (Table 5.b). This proves that this can be taken as a selection variable. Indeed, in a separate probit equation (not shown here), this variable results to be not influential on overeducation, which is consistent with the requirements for a selection variable. The coefficient *rho*, measuring the potential bias in the sample selection, has a positive sign but it is not statistically significant.

TABLES 4a AND 4.b ABOUT HERE

Finally, Table A6.a reports the estimate of a model concerning workers with a university degree where the dummy for the movers has been replaced with the distance in kilometres (and its square) between the province where the individual lives in the present and the province where he/she comes from. This model confirms all the results of the previous estimation. Furthermore, it reveals that both the kilometers and its square are strongly statistically significant. The sign of their coefficients points out that the risk of overeducation decreases at a decreasing rate with the migration distance.

Even in this case the selection variable passes the test, as it is strongly correlated with employment (Table A6.b) but not with overeducation (not shown here). As in previous estimates, the *rho* coefficient is positive but not significant.

[6] Note that in German local labour markets tenure has a strong negative effect on overeducation risk, whereas age affects participation but does not affect overeducation (Büchel & Van Ham, 2003).

5. Concluding remarks

Among the various theories which have been advanced to explain overeducation, the goal of this paper was to test the relevance of the so called spatial flexibility hypothesis. According to it, overeducation arises when workers cannot afford the costs of, or dislike, mobility needed to search for a suitable job if demand and supply of skills do not have a homogeneous spatial distribution. 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 poor learning due to a low quality of the services provided by the educational institutions.

As a measure of individual spatial flexibility, our dataset allowed us to consider the time spent in commuting by the worker and if he/she has moved for working reasons related to his/her current job, as well as the distance in kilometres between the province where the individual lives in the present and the province where he/she comes from. To the best of our knowledge these measures of mobility represent an innovation with respect to previous literature on this topic. Besides these two factors, we included in the model a set of variables characterising the LLM where the worker resides. Furthermore, this paper represents the first attempt to evaluate the influence of spatial mobility on overeducation on Italian data.

After testing for the possible endogeneity of the migration choice, and in order to overcome the problem of the selection bias due to the correlation between employment and overeducation, we estimated a bivariate probit model with sample selection. Our most important finding is that spatial flexibility matters. According to our results, as for workers holding an upper secondary degree, the risk of overeducation decreases with commuting time, while, among the university graduates, movers are less overeducated than stayers. Furthermore, a longer migration distance decreases overeducation risks. These results imply that frictions and barriers increasing the costs of spatial mobility worsen the matching between required and possessed education in the labour market.

Such evidence suggests that a more efficient valorisation of human capital could take advantage of a reduction of both monetary and non monetary costs of labour spatial mobility. This could require proper transportation facilities and services, flexibility of working time and the availability of complementary services, as child care. Moreover, favourable conditions of the housing market represent an important factor to boost mobility. Finally, 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.

References

Audas, R., & Dolton P. (1999). The Effect of Local Labour Market Conditions on the Decision to Migrate Among UK Graduates. Paper presented at the Royal Economic Society Conference, Nottingham.

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.

Brunello, G., Comi, S., & Lucifora C. (2001). The Returns to Education in Italy. In C. Harmon, I. Walker and N. Westergard-Nielsen (eds.), *Education and Earnings in Europe: A Cross Country Analysis of the Returns to Education*. Edward Elgar, Cheltenham, UK.

Büchel, F., & Battu, H. (2003). The Theory of Differential Overqualification: Does It Works? *Scottish Journal of Political Economy*, 50 (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. Paper presented to the 16th SASE Annual Meeting on Socio-Economics, George Washington University, Washington D.C. USA, July 8-11, 2004.

Cannari, L., Nucci, F., & Sestito, P. (2000). Geographical Labour Mobility and the Cost of Housing: Evidence from Italy. *Applied Economics*, 32, 1899-1906.

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

Croce, G., & Ghignoni, E. (2010). Employer-provided training and knowledge spillovers. Evidence from Italian local labour markets. *Regional Studies*, dx.doi.org/10.1080/00343404.2010.497136.

Dalmazzo, A., & De Blasio, G. (2007). Social Returns to Education in Italian Local Labour Markets. *Annals of Regional Sciences*, 41, 51–69.

Davidson, J., & MacKinnon, J. (1993). *Estimation and Inference in Econometrics*. Oxford University Press, Oxford.

Di Addario, S., & Patacchini, E. (2008). Wages and the City. Evidence from Italy. *Labour Economics*, 15, 1040–1061.

Di Pietro, G., & Cutillo, A. (2006). University Quality and Labour Market Outcomes in Italy. *Labour*, 20(1), 37–62.

Di Pietro, G., & Urwin, P. (2006). Education and Skills Mismatch in the Italian Graduate Labour Market. *Applied Economics*, 38, 79–93.

Dostie, B., & Leger, P.T. (2009). Self-Selection in Migration and Returns to Unobservable Skills. *Journal of Population Economics*, 22(4), 1005–1024.

Duncan, G., & Hoffman, S. (1981). The Incidence and Wage Effects of Overeducation. *Economics of Education Review*, 1(1), 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. (2011). Few and underutilized? Overeducation of Italian graduates, in E. Mandrone (ed.), *PLUS Empirical Papers*, Temi e Ricerche, Isfol Editore, in press.

Hartog, J. (2000). Over-Education and Earnings: Where Are We, Where Should We Go? *Economics of Education Review*, 19, 131–147.

Hartog, J., & Oosterbeek, H. (1988). Education, Allocation and Earnings in the Netherlands: Overschooling? *Economics of Education Review*, 7(2), 185–194.

Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251–1261.

Hensen, M., & de Vries, R. (2004). The Relationship between Geographical Mobility and Education-Job Mismatches. ERSA Conference Papers Ersa04p158, European Regional Science Association.

Istat (1997). *I sistemi locali del lavoro 1991*. Collana Argomenti, 10, Rome.

Istat (2003). *I laureati e il mercato del lavoro, Inserimento professionale dei laureati - Indagine 2001*, Rome.

Istat (2006). *I laureati e il mercato del lavoro, Inserimento professionale dei laureati - Indagine 2004*, Rome.

Istat (2010). *I laureati e il mercato del lavoro, Inserimento professionale dei laureati - Indagine 2007*, Rome.

Kiker, B., Santos, M., & Mendes de Oliveira, M. (1997). Overeducation and Undereducation: Evidence for Portugal. *Economics of Education Review*, 16(2), 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, 939, June.

Linsley, I. (2005b). Causes of Overeducation in the Australian Labour Market. Research Paper, University of Melbourne, Department of Economics, Melbourne Victoria, Australia, 940, June.

McGoldrick, K.M. & Robst, J. (1996). Gender Differences in Overeducation: A Test of the Theory of Differential Overqualification. *American Economic Review*, 86(2), 280–284.

McGuinness, S. (2006). Overeducation in the Labour Market. *Journal of Economic Surveys*, 20(3), 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*, 72(2), 334–338.

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

Ordine, P., & Rose, G. (2009b). Inefficient Self-Selection into Education and Wage Inequality in Italy. Paper presented to the XXIV National Conference of Labour Economics, University of Sassari, 24-25 September.

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), 24–50.

Sanroma, E., & Ramos, R. (2004). Overeducation and Local Labour Markets in Spain. Paper presented to the 44th European Regional Science Association Congress, Porto, 25-29 August.

Sattinger, M. (1993). Assignment Models of the Distribution of Earnings. *Journal of Economic Literature*, 31, 831–880.

Sloane, P.J. (2003). Much Ado About Nothing? What Does the Overeducation Literature Really Tell Us? In F. Büchel, A. De Grip, A. Mertens (Eds.), *Overeducation in Europe. Current Issues in Theory and Policy*, Edward Elgar, Cheltenham, UK, Northampton MA, USA, 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*, 17, 229–252.

Van Ham, M. (2003). *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), 690–695.

Table 1
Descriptive statistics

Variable description	Employed			Full sample		
	N. Obs.	Mean	Std. Dev.	N. Obs.	Mean	Std. Dev.
Dependent variable						
Overeducated = 1	12,694	0.333	0.471			
Employed = 1				19335	0.657	0.475
Individual characteristics						
Age	12,694	39.492	12.148	19,335	37.726	11.466
Age squared	12,694	1707.22	993.028	19,335	1554.71	920.759
Specific experience	12,694	11.991	10.969	12,694	11.991	10.969
Specific experience squared	12,694	264.077	371.881	12,694	264.077	371.881
Dependent employee, permanent contract	12,694	0.668	0.471	12,694	0.668	0.471
Dependent employee, fixed term contract	12,694	0.148	0.355	12,694	0.148	0.355
Self employed	12,694	0.184	0.387	12,694	0.184	0.387
Foreigner	12,694	0.011	0.104	19,335	0.014	0.118
Children 0-12	12,694	0.779	0.415	19,335	0.816	0.387
High marks	12,694	0.509	0.500	19,335	0.479	0.500
Regular educational path	12,694	0.777	0.416	19,335	0.768	0.422
Internet	12,694	0.854	0.353	19,335	0.809	0.393
English	12,694	0.354	0.478	19,335	0.342	0.474
Daily commuting time (home to work)	12,694	20.806	20.197	10,903	20.806	20.197
Movers	12,694	0.123	0.328	12,694	0.123	0.328
Upper secondary education	12,694	0.667	0.471	19,335	0.711	0.453
Tertiary education	12,694	0.312	0.463	19,335	0.271	0.445
Post tertiary education	12,694	0.022	0.146	19,335	0.017	0.130
Municipality size: less than 10,000 inhabitants	12,694	0.272	0.445	19,335	0.284	0.451
Municipality size: 10,001 – 100,000 inhabitants	12,694	0.358	0.479	19,335	0.375	0.484
Municipality size: more than 100,000 inhabitants	12,694	0.371	0.483	19,335	0.341	0.474
Males	12,694	0.465	0.499	19,335	0.382	0.486
Females	12,694	0.535	0.499	19,335	0.618	0.486
Divorced/Separated/Widow/Single	12,694	0.436	0.496	19,335	0.433	0.495
Married/Cohabitant	12,694	0.564	0.496	19,335	0.567	0.495
Father education: max lower secondary	11,985	0.678	0.467	18,201	0.692	0.462
Father education: Upper secondary	11,985	0.232	0.422	18,201	0.225	0.418
Father education: Tertiary degree	11,985	0.090	0.286	18,201	0.083	0.276
North West	12,694	0.240	0.427	19,335	0.213	0.410
North East	12,694	0.214	0.410	19,335	0.189	0.391
Centre	12,694	0.207	0.405	19,335	0.203	0.402
South	12,694	0.238	0.426	19,335	0.276	0.447
Islands	12,694	0.100	0.301	19,335	0.118	0.323
Firms characteristics						
Production of goods	12,694	0.146	0.353	12,562	0.146	0.353
Production services	12,562	0.192	0.394	12,562	0.192	0.394
Distribution services	12,562	0.056	0.230	12,562	0.056	0.230
Personal services	12,562	0.407	0.491	12,562	0.407	0.491
Social services	12,562	0.199	0.400	12,562	0.199	0.400
Small firm (10-49 employees)	7,839	0.707	0.455	7,839	0.707	0.455
Medium firm (50-499 employees)	7,839	0.215	0.411	7,839	0.215	0.411
Big firm (500 or more employees)	7,839	0.077	0.267	7,839	0.077	0.267
Public sector	12,647	0.383	0.486	12,647	0.383	0.486
LLM characteristics						
Mountain area (km2)	12,694	320.053	346.864	19,335	307.900	345.640
University centre	12,694	0.541	0.498	19,335	0.516	0.500
% Foreigner	12,694	4.665	2.621	19,335	4.366	2.652
Average firm size	12,694	8.005	2.898	19,335	7.737	2.963
Industrial district	12,694	0.180	0.384	19,335	0.169	0.375
(log) Employment density	12,694	4.811	1.251	19,335	4.766	1.273
% Upper Secondary and Tertiary graduates	12,694	34.861	5.316	19,335	34.460	5.480
Accident/1000 cars	12,694	7.385	3.403	19,335	7.100	3.405
Local unemployment rate	12,694	5.319	3.706	19,335	5.831	3.904
Selection variable						
Loan for accommodation	12,694	0.189	0.392	19,335	0.174	0.379

Table 2
Incidence of overeducation by educational degree

Variable description	Upper secondary education	Tertiary education	Total
Individual characteristics			
20-29 years	48.07	23.50	42.21
30-39 years	39.87	24.40	34.10
40-49 years	39.77	14.96	31.88
50-64 years	32.87	12.29	24.62
Specific experience 1-4 years	49.30	21.94	40.17
Specific experience 5-9 years	42.98	26.69	38.18
Specific experience 10-15 years	39.94	16.13	32.45
Specific experience 16 and more	29.68	11.39	23.06
Dependent employee, permanent contract	36.59	17.54	30.74
Dependent employee, fixed term contract	50.93	19.48	39.33
Self employed	49.46	19.70	37.57
Italian	40.32	17.96	32.86
Foreigner	78.49	53.33	70.29
No children 0-12	41.48	19.32	32.42
Children 0-12	40.55	17.97	33.51
Low marks	43.88	22.09	41.14
High marks	35.04	17.48	25.66
Irregular educational path	44.09	26.91	41.79
Regular educational path	39.37	17.49	30.82
No Internet	48.74	15.12	41.89
Internet	39.05	18.65	31.80
No English	40.77	15.9	34.28
English	40.65	20.83	31.43
Daily commuting time: less than 30 minutes	40.11	18.39	33.12
Daily commuting time: 31 minutes and more	34.84	17.93	27.87
No movers	41.68	19.41	34.54
Movers	32.74	12.56	24.16
Municipality size: less than 10,000 inhabitants	40.27	20.19	35.45
Municipality size: 10,001 – 100,000 inhabitants	39.37	16.28	32.31
Municipality size: more than 100,001 inhabitants	42.80	18.99	32.59
Males	41.89	17.85	34.63
Females	39.64	18.69	32.08
Divorced/Separated/Widow/Single	45.58	20.84	37.77
Married/Cohabitant	36.81	16.58	29.79
Father education: at most lower secondary	40.12	18.68	34.80
Father education: upper secondary	40.52	18.20	30.32
Father education: tertiary education	40.96	17.32	23.75
North West	42.57	18.68	34.97
North East	39.32	21.48	33.91
Centre	43.94	20.54	36.11
South	36.77	15.02	28.97
Islands	41.90	15.64	32.16
Firms characteristics			
Production of goods	45.83	28.92	42.77
Production services	35.16	24.54	31.09
Distribution services	53.10	47.86	52.34
Personal services	64.31	40.98	60.26
Social services	22.93	9.44	16.16
Firm size: small (10-49)	48.50	25.38	42.79
Firm size: medium (50-499)	47.20	30.70	43.12
Firm size: big (500 or more)	44.47	31.32	40.53
Public sector	24.45	10.74	18.06
Private sector	47.96	27.15	42.75
Tot.	40.74	18.34	33.27

Table 3
Probability of overeducation (without sample selection); movers and daily commuting time

Variables	Total sample		Upper secondary graduates		Tertiary graduates	
	Coef.	z	Coef.	z	Coef.	z
Individual characteristics						
Age	0.0473**	3.75	0.0495**	3.66	0.0311	1.01
Age squared	-0.0006**	-4.01	-0.0006**	-3.88	-0.0004	-1.23
Female	-0.0261	-0.82	-0.0821*	-2.47	0.0779	1.21
Married	-0.0321	-0.87	-0.0532	-1.27	0.0474	0.74
Specific experience	-0.0189**	-3.59	-0.0209**	-2.86	-0.0118	-1.04
Specific experience squared	0.0002	1.71	0.0003	1.62	0.0001	0.24
Foreigner	0.6661**	6.43	0.6926**	4.67	0.5018*	1.96
Children 0-12	-0.0489	-1.42	-0.0230	-0.52	-0.0811	-1.26
High marks	-0.2139**	-7.65	-0.2355**	-7.40	-0.1057	-1.56
Regular path	-0.1162**	-3.77	-0.0844*	-2.16	-0.2683**	-2.86
Internet	-0.2851**	-7.59	-0.3318**	-7.84	-0.0839	-0.89
English	-0.0533	-1.49	-0.1242**	-3.19	0.0087	0.17
Technical skills	-	-	-0.3654**	-6.93	-	-
Daily commuting time (home to work)	-0.0020**	-2.81	-0.0027**	-3.14	-0.0003	-0.23
Movers	-0.1471**	-2.97	-0.0877	-1.45	-0.2893**	-3.71
Education (ref. upper secondary education)						
Tertiary education	-0.2020**	-4.55	-	-	-	-
Post Tertiary education	-0.3885**	-3.04	-	-	-	-
Employment contract (ref. permanent contract)						
Dependent employee, fixed term contract	0.0467	1.16	0.1262**	2.65	-0.2194*	-2.17
Self employed	-0.0140	-0.27	0.0347	0.66	-0.2197*	-2.12
Father education (ref. less than primary)						
Primary education	-0.1256	-1.55	-0.1037	-1.16	-0.2353	-1.19
Lower secondary education	-0.1358	-1.64	-0.1261	-1.39	-0.2269	-1.14
Upper secondary education	-0.2303*	-2.54	-0.2148*	-2.17	-0.4162*	-2.17
Tertiary education	-0.2320*	-2.02	-0.1848	-1.38	-0.4171*	-1.96
Firm characteristics						
Firm size (ref. small firm: 10-49)						
Medium firm (50-499)	0.1153*	2.36	0.0849	1.36	0.1253	1.51
Big firm (501 or more)	0.1288	1.73	0.0956	1.07	0.0981	0.84
Public sector	-0.1312*	-2.47	-0.2383**	-4.01	0.0440	0.51
Constant	-0.1533	-0.67	0.2542	1.00	-0.2169	-0.36
Other controls						
Mother education	yes		yes		yes	
Branch of economic activity	yes		yes		yes	
Macroarea	yes		yes		yes	
Number of obs.	10,128		6,653		3,475	
Pseudo R2	0.1361		0.1054		0.1435	

Notes: Cluster adjusted robust standard errors. ** 1% significance level; * 5% significance level.

Table 4.a

Upper secondary graduates: overeducation equation with sample selection (selection variable: weighted loan for accommodation)

Variables	Coef.	Robust Std. Err.	z	P>|z|
Individual characteristics				
Age	0.0427	0.026	1.66	0.098
Age squared	-0.0006	0.000	-1.99	0.047
Specific experience	-0.0161	0.010	-1.56	0.118
Specific experience squared	0.0002	0.000	0.55	0.584
Foreigner	0.9932	0.305	3.26	0.001
Children 0-12	-0.0115	0.060	-0.19	0.847
High marks	-0.1887	0.042	-4.54	0.000
Regular path	-0.0965	0.055	-1.75	0.080
Internet	-0.3308	0.064	-5.18	0.000
English	-0.1931	0.058	-3.32	0.001
Technical skills	-0.2304	0.068	-3.37	0.001
Daily commuting time (home to work)	-0.0034	0.001	-2.73	0.006
Movers	-0.0257	0.080	-0.32	0.747
Employment contract (ref. permanent contract)				
Dependent employee, fixed term contract	0.1236	0.090	1.38	0.169
Self employed	0.0642	0.070	0.92	0.356
Father education (ref. less than primary)				
Primary education	-0.1918	0.100	-1.91	0.056
Lower secondary education	-0.1536	0.110	-1.40	0.162
Upper secondary education	-0.2373	0.117	-2.02	0.043
Tertiary education	-0.2309	0.203	-1.14	0.255
Municipality size (ref. 0-10,000)				
Inhabitants: 10,001 to 100,000	-0.0893	0.065	-1.38	0.168
Inhabitants: 100,001 and more	0.0576	0.115	0.50	0.618
Firm characteristics				
Firm size (ref. small firm: 10-49)				
Medium firm (50-499)	0.0698	0.072	0.97	0.332
Big firm (501 or more)	0.0373	0.101	0.37	0.711
Public sector	-0.1531	0.087	-1.77	0.077
Local variables				
Mountain area (km2)	-0.0001	0.000	-0.99	0.324
University centre	0.0790	0.077	1.03	0.304
% Foreigner	-0.0161	0.016	-0.98	0.326
Average firm size	0.0146	0.013	1.12	0.264
Industrial district	-0.1436	0.067	-2.13	0.033
(log) Employment density	-0.0095	0.036	-0.26	0.794
% Upper Secondary graduates	0.0171	0.011	1.50	0.134
Accident/1000 cars	0.0011	0.012	0.09	0.928
Local unemployment rate	-0.0183	0.017	-1.05	0.294
Constant	-0.3429	0.638	-0.54	0.591
Other controls				
Interactions between city/rural area, gender and married/single	yes			
Mother education	yes			
Branch of economic activity	yes			
Macroarea	yes			
Number of obs.	3,658			

Cluster adjusted robust standard errors

Table 4.b

Upper secondary graduates: employment equation (selection variable: weighted loan for accommodation)

Variables	Coef.	Robust Std. Err.	z	P>|z|
Individual characteristics				
Age	-0.1488	0.011	-13.21	0.000
Age squared	0.0021	0.000	16.27	0.000
Foreigner	-0.3562	0.108	-3.31	0.001
Children 0-12	-0.1640	0.033	-4.95	0.000
High marks	0.1292	0.026	5.06	0.000
Regular path	0.0874	0.028	3.12	0.002
Internet	0.5548	0.028	20.13	0.000
English	-0.0652	0.028	-2.32	0.020
Technical skills	0.1549	0.037	4.19	0.000
Father education (ref. less than primary)				
Primary education	0.0492	0.062	0.80	0.425
Lower secondary education	-0.0444	0.066	-0.67	0.504
Upper secondary education	-0.0472	0.073	-0.64	0.519
Tertiary education	-0.1754	0.113	-1.55	0.122
Municipality size (ref. 0-10,000)				
Inhabitants: 10,001 to 100,000	0.0712	0.050	1.41	0.158
Inhabitants:100,001 and more	0.1151	0.165	0.70	0.484
Local variables				
Mountain area (km2)	0.0001	0.000	0.37	0.714
University centre	0.0557	0.135	0.41	0.681
% Foreigner	-0.0087	0.040	-0.22	0.828
Average firm size	0.0166	0.018	0.91	0.360
Industrial district	0.0889	0.136	0.65	0.514
(log) Employment density	-0.0665	0.073	-0.91	0.365
% Upper Secondary graduates	-0.0076	0.023	-0.32	0.747
Accident/1000 cars	-0.0028	0.028	-0.10	0.920
Local unemployment rate	-0.0767	0.037	-2.09	0.037
Selection variable				
Weighted loan for accommodation	0.0191	0.000	41.28	0.000
Constant	3.1203	0.602	5.18	0.000
Other controls				
Interactions between city/rural area, gender and married/single	yes			
Mother education	yes			
Macroarea	yes			
Number of obs.	6,583			
/athrho	-0.1168	0.114	-1.02	0.307
rho	-0.1163	0.113		
Wald test of independent equations (rho = 0): chi2(1) = 1.04		Prob > chi2 = 0.3073		

Cluster adjusted robust standard errors

Table 5.a
Tertiary graduates: overeducation equation with sample selection (selection variable: loan for accommodation)

Variables	Coef.	Robust Std. Err.	z	P>|z|
Individual characteristics				
Age	0.0216	0.030	0.70	0.477
Age squared	-0.0004	0.000	-0.90	0.325
Specific experience	-0.0104	0.012	-0.90	0.368
Specific experience squared	0.0000	0.000	0.10	0.924
Foreigner	0.5654	0.258	2.10	0.028
Children 0-12	-0.0708	0.070	-1.00	0.314
High marks	-0.1232	0.071	-1.70	0.084
Regular path	-0.2578	0.101	-2.50	0.011
Internet	-0.0951	0.095	-1.00	0.317
English	0.0008	0.052	0.00	0.987
Daily commuting time (home to work)	-0.0005	0.001	-0.30	0.708
Movers	-0.2983	0.079	-3.70	0.000
Employment contract (ref. permanent contract)				
Dependent employee, fixed term contract	-0.2146	0.101	-2.10	0.033
Self employed	-0.2165	0.101	-2.10	0.033
Father education (ref. less than primary)				
Primary education	-0.2651	0.198	-1.30	0.181
Lower secondary education	-0.2516	0.199	-1.20	0.205
Upper secondary education	-0.4411	0.194	-2.20	0.023
Tertiary education	-0.4497	0.213	-2.10	0.034
Municipality size (ref. 0-10,000)				
Inhabitants: 10,001 to 100,000	-0.0728	0.088	-0.80	0.409
Inhabitants:100,001 and more	-0.0572	0.113	-0.50	0.611
Firm characteristics				
Firm size (ref. small firm: 10-49)				
Medium firm (50-499)	0.1406	0.081	1.70	0.082
Big firm (501 or more)	0.1094	0.114	0.90	0.339
Public sector	0.0573	0.090	0.60	0.522
Local variables				
Mountain area (km2)	0.0001	0.000	0.70	0.441
University centre	0.1813	0.102	1.70	0.076
% Foreigner	-0.0198	0.022	-0.80	0.373
Average firm size	0.0016	0.012	0.10	0.887
Industrial district	-0.0995	0.097	-1.00	0.303
(log) Employment density	-0.0412	0.050	-0.80	0.405
% Tertiary graduates	-0.0509	0.024	-2.10	0.034
Accident/1000 cars	0.0320	0.013	2.40	0.014
Local unemployment rate	-0.0219	0.019	-1.10	0.257
Constant	0.5472	0.739	0.70	0.459
Other controls				
Interactions between city/rural area, gender and married/single	yes			
Mother education	yes			
Branch of economic activity	yes			
Macroarea	yes			
Number of obs.	3,475			

Cluster adjusted robust standard errors

Table 5.b
Tertiary graduates: employment equation (selection variable: loan for accommodation)

Variables	Coef.	Robust Std. Err.	z	P>|z|
Individual characteristics				
Age	0.0219	0.014	1.50	0.129
Age squared	0.0004	0.000	2.40	0.014
Foreigner	-0.9113	0.128	-7.10	0.000
Children 0-12	-0.3916	0.026	-14.80	0.000
High marks	0.1724	0.039	4.40	0.000
Regular path	0.4819	0.053	9.10	0.000
Internet	0.3411	0.045	7.50	0.000
English	-0.0148	0.026	-0.50	0.572
Father education (ref. less than primary)				
Primary education	-0.1248	0.098	-1.20	0.201
Lower secondary education	-0.1017	0.100	-1.00	0.311
Upper secondary education	-0.1380	0.103	-1.30	0.180
Tertiary education	-0.2301	0.116	-1.90	0.048
Municipality size (ref. 0-10,000)				
Inhabitants: 10,001 to 100,000	-0.0169	0.062	-0.20	0.784
Inhabitants:100,001 and more	-0.2033	0.197	-1.00	0.303
Local variables				
Mountain area (km2)	0.0001	0.000	0.40	0.635
University centre	0.1431	0.204	0.70	0.483
% Foreigner	0.0365	0.050	0.70	0.462
Average firm size	0.0067	0.023	0.20	0.772
Industrial district	0.0203	0.155	0.10	0.896
(log) Employment density	-0.0404	0.088	-0.40	0.648
% Tertiary graduates	-0.0208	0.059	-0.30	0.725
Accident/1000 cars	0.0001	0.031	0.00	0.998
Local unemployment rate	-0.0184	0.041	-0.40	0.649
Selection variable				
Loan for accommodation	0.3303	0.034	9.80	0.000
Constant	-0.9514	0.582	-1.60	0.102
Other controls				
Interactions between city/rural area, gender and married/single	yes			
Mother education	yes			
Macroarea	yes			
Number of obs.	4,762			
/athrho	0.0064	0.135	0.005	0.962
rho	0.0064	0.135		
Wald test of independent equations (rho = 0): chi2(1) = 0.00		Prob > chi2 = 0.9622		

Cluster adjusted robust standard errors

Appendix A

Table A6.a

Tertiary graduates: overeducation equation with sample selection; Km for movers (selection variable: loan for accommodation)

Variables	Coef.	Robust Std. Err.	z	P>|z|
Individual characteristics				
Age	0.0215	0.031	0.70	0.482
Age squared	-0.0003	0.000	-0.94	0.348
Specific experience	-0.0099	0.012	-0.85	0.398
Specific experience squared	0.0000	0.000	0.04	0.967
Female	0.0487	0.065	0.75	0.456
Married	0.0736	0.065	1.14	0.255
Foreigner	0.5039	0.253	2.00	0.046
Children 0-12	-0.0717	0.068	-1.05	0.295
High marks	-0.1187	0.069	-1.71	0.087
Regular path	-0.2471	0.099	-2.49	0.013
Internet	-0.0848	0.094	-0.91	0.364
English	-0.0036	0.052	-0.07	0.944
Daily commuting time (home to work)	-0.0005	0.001	-0.37	0.712
Km for movers	-0.0016	0.001	-3.23	0.001
Km^2 for movers	0.0001	0.000	2.52	0.012
Employment contract (ref. permanent contract)				
Dependent employee, fixed term contract	-0.2128	0.101	-2.10	0.035
Self employed	-0.2073	0.102	-2.03	0.042
Father education (ref. less than primary)				
Primary education	-0.2867	0.198	-1.45	0.148
Lower secondary education	-0.2782	0.199	-1.40	0.162
Upper secondary education	-0.4612	0.192	-2.41	0.016
Tertiary education	-0.4721	0.213	-2.22	0.027
Municipality size (ref. 0-10,000)				
Inhabitants: 10,001 to 100,000	-0.0782	0.087	-0.89	0.371
Inhabitants:100,001 and more	0.0433	0.122	0.36	0.722
Firm characteristics				
Firm size (ref. small firm: 10-49)				
Medium firm (50-499)	0.1391	0.082	1.71	0.088
Big firm (501 or more)	0.1123	0.114	0.99	0.324
Public sector	0.0590	0.092	0.64	0.520
Local variables				
Mountain area (km2)	0.0001	0.000	1.12	0.265
University centre	0.1915	0.099	1.93	0.053
% Foreigner	-0.0167	0.021	-0.78	0.437
Average firm size	0.0021	0.012	0.18	0.854
Industrial district	-0.1015	0.097	-1.05	0.296
(log) Employment density	-0.0303	0.050	-0.61	0.544
% Tertiary graduates	-0.0531	0.026	-2.07	0.038
Accident/1000 cars	0.0316	0.013	2.44	0.015
Local unemployment rate	-0.0228	0.020	-1.16	0.247
Constant	0.3518	0.715	0.49	0.623
Other controls				
Interactions between city/rural area, gender and married/single	no			
Mother education	yes			
Branch of economic activity	yes			
Macroarea	yes			
Number of obs.	3,475			

Cluster adjusted robust standard errors

Table A6.b

Tertiary graduates: employment equation; Km for movers (selection variable: loan for accommodation)

Variables	Coef.	Robust Std. Err.	z	P>|z|
Individual characteristics				
Age	0.0103	0.015	0.70	0.486
Age squared	0.0006	0.000	3.46	0.001
Female	-0.2037	0.030	-6.89	0.000
Married	0.0691	0.033	2.11	0.035
Foreigner	-0.9027	0.130	-6.93	0.000
Children 0-12	-0.3767	0.025	-14.87	0.000
High marks	0.1889	0.039	4.90	0.000
Regular path	0.4511	0.053	8.50	0.000
Internet	0.3899	0.044	8.91	0.000
English	-0.0101	0.025	-0.39	0.693
Father education (ref. less than primary)				
Primary education	-0.1457	0.093	-1.56	0.118
Lower secondary education	-0.1195	0.096	-1.25	0.212
Upper secondary education	-0.1669	0.099	-1.69	0.091
Tertiary education	-0.2508	0.111	-2.26	0.024
Municipality size (ref. 0-10,000)				
Inhabitants: 10,001 to 100,000	-0.0264	0.064	-0.41	0.679
Inhabitants:100,001 and more	0.0690	0.115	0.60	0.549
Local variables				
Mountain area (km2)	0.0001	0.000	0.67	0.502
University centre	0.1570	0.204	0.77	0.443
% Foreigner	0.0331	0.050	0.67	0.506
Average firm size	0.0067	0.023	0.29	0.773
Industrial district	0.0026	0.156	0.02	0.987
(log) Employment density	-0.0201	0.089	-0.23	0.821
% Tertiary graduates	-0.0249	0.061	-0.41	0.682
Accident/1000 cars	0.0015	0.032	0.05	0.962
Local unemployment rate	-0.0167	0.041	-0.41	0.683
Selection variable				
Loan for accommodation	0.3181	0.033	9.53	0.000
Constant	-0.9187	0.611	-1.50	0.133
Other controls				
Interactions between city/rural area, gender and married/single	no			
Mother education	yes			
Macroarea	yes			
Number of obs.	4,762			
/athrho	0.0813	0.121	0.67	0.502
rho	0.0811	0.120		
Wald test of independent equations (rho = 0): chi2(1) = 0.45		Prob > chi2 = 0.5016		

Cluster adjusted robust standard errors

Appendix B
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)

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

Technical skills: Isfol Plus 2005, question V882 (Type of Upper secondary degree)

Employment contract (ref. cat. : dependent employee with permanent contract): Isfol Plus 2005, question V110_2

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)

Children 0-12: 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?)

Kilometers: road distance between Province of origin and Province of actual residence. Isfol Plus 2005, question V440 (Which Province do you come from?)

Father education: Isfol Plus 2005, dummy variable built on question V1050_1

Mother education: Isfol Plus 2005, dummy variable built on question V1050_2

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).

Public sector: Isfol Plus 2005, question V390.

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 (by LLMs, when not differently specified):

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.
Unemployment rate of the province of origin: *Censimento della Popolazione* 1981, 1991, 2001 (average data), Istat.

Selection variable:

(Weighted) Loan for accommodation: Isfol Plus 2005, dummy variable built on questions V1670_1 (Do you pay a loan for your accommodation?). This variable has been weighted by using question V1080_1: what share of total household income is represented by your individual income?