Does it pay studying far from home? Explaining the returns to geographic mobility of Italian graduates*

Mattia Makovec†
Università Bocconi
European Centre for Social Welfare Policy and Reasearch
mattia.makovec@uni-bocconi.it

Very preliminary and incomplete

February 14, 2005

Abstract

The main idea of this paper is that in a labour market with high regional imbalances, like the Italian one, early occupational earnings of young graduates vary significantly not only according to the region of work, but also according to the region where university was attended. This issue has been largely ignored by the current empirical research on regional wage differentials. Using a propensity-score matching estimation approach, I show the existence of a positive and significant wage premium associated to attending university in the North rather than in the South. The identification of the appropriate treatment and control groups is allowed by the detailed information about residence before university, during university and during work available in the data, a representative sample of the 1998 national cohort of graduates surveyed by the National Statistical Institute (ISTAT) three years after the conclusion of their studies. JEL Code: C21, C31, I21, J24, J31, J61, R23.

*Acknowledgements: I thank Giovanni Seri and Luisa Franconi for making the data available at Laboratorio Adele, ISTAT and for their helpful assistance. I benefitted from comments and suggestions by the participants to the "First Italian Congress of Econometrics and Empirical Economics", in particular Giacomo De Giorgi and Michele Pellizzari. I am very much in debt with Daniela Feliziani for giving me the opportunity to carry out this research. Financial support by Fondazione Cariplo is greatefully acknowledged. The usual disclaimer applies.

†Address for correspondence: Mattia Makovec, European Centre for Social Welfare Policy and Research, Berggasse 17, 1090, Vienna, Austria. Email: makovec@euro.centre.org.

1 Introduction and motivation

A bunch of empirical studies have been dedicated to the analysis of regional wage differentials in Italy, mainly in the context of estimating national earnings equations, reporting substantial wage differentials across regions. The approach based on standard Mincerian regression functions, though, postulate some strong assumptions on the rates of return to education across the population, which are normally assumed to be constant (or homogenous) by levels of educational attainment. Regional dummies included in linear wage regressions are normally assumed to capture the premium of working in a given geographic area of the country with respect to another, after controlling for a number of observables; these dummies however, are most likely endogenous, an aspect which is usually not tackled by the literature. Other techniques adopted to measure regional wage differentials include decomposition methods and nonparametric methods to compare the wage distributions across different regions.

The approach followed in this paper is quite different. First of all, the focus is on a specific fraction of the population: the graduates at their earlier stages of their labour market career. Further, convincing evidence is provided on the following hypothesis. For young graduates, the wage premium associated to working in the North rather than in the South, is the sum of two effects: the first is purely associated to the geographic location of the job place, while the second is associated to the geographic area where university were attended. In other words, in a labour market highly segmented at the regional level, like the Italian one, not only where people work, but also where people study matters for their early occupational earnings. The first effect is identified comparing the wages of students coming from the same geographic area, who studied in the same region but ended up working in different regions ("job place effect"); the second effect is identified comparing the wages of students coming from the same geographic area, who ended up working in the same region, but after attending university in different regions ("university location effect").

2 The data

In this study, I use a valuable data source based on a survey carried out by the National Statistical Institute (ISTAT) on the labour market outcomes of a representative sample of Italian former university students. The dataset reports monthly net wages and contains a wealth of information on labour market characteristics and histories, demographics, and family background of the cohort of 1998

graduates interviewed in 2001. Some basic descriptive statics of the sample are reported in table 1.

Table 1 here

The survey has been carried out by ISTAT for several years starting from the mid '80s, usually with intervals of 3 years between each survey. Individuals are sampled first at the university level, among the population of university students. The questionnaire contains detailed information of individual characteristics both before and after university enrollment, namely family background (father and mother education and occupational status), region (and province) of residence, region (and province) of study during university, a series of indicators of performance, both at the university and high school level (both university and high school leaving grade), together with a wealth of information on the labour market career of the students after graduation, unemployment experience, occupational status, profession and sector of economic activity.

The original sample includes 20844 individuals randomly sampled among the total graduate population. The sample dimension is considerable, since the ratio between sampled person and the universe is roughly 1:5.

3 Geographic (im)mobility of Italian graduates

Table 2 and table 3 provide some evidence on the geographical mobility of Italian graduates. The detailed information available in the dataset, in fact, enables identifying the province (and therefore the region and the macro-region) of residence of the student before university enrolment, during the attendance of university, and at the end of university, if they are employed.

Table 2 here

The tables show a high degree of immobility in all macro-regions considered (North, Centre, South), both from pre-university to university residence, and from residence during university to residence of work. In other words, graduate students tend to both study and work in their region of origin. In particular, the region of study during university has great influence on the choice of the region of work. This is evident looking at students coming from Southern Italy: for those who migrate to the North during their university studies, it is easier to find a job in the North, while among the residents from the South, who attended also a university in the South, only about 10% will be able to work in the North.

The low degree of mobility towards employment in Northern regions of graduates from universities in the South, might be interpreted as one of the causes at the origin of regional wage imbalances: having studied in a university from the South might be gauged as a "stigma" by the employers in the North, so that university students are stuck searching for a job in the South.

Table 3 here

4 Estimation and results

The bunch of detailed information available in the dataset at geographic level, enables the identification of the the two effects of interest. Knowing the place of residence of the students before, during and after university attendance allows building up appropriate treatment and controls groups for both effects. Table 4 illustrates the two population subgroups in the case of students resident in the South before university enrolment.

Table 4 here

In particular, the wage premium due to the "Job place effect" is defined as the premium associated to working in the North for students coming from the South and studying in the South, relative to students coming from, studying and working in the South. While the wage premium due to the "University location effect" in this case is defined as the premium associated to studying in the North for students coming from the South and working in the North, relative to students coming from the South, studying in the South and working in the North.

More formally, both effects can be considered as two "Average Treatment on the Treated" effects (ATT) in the language of the program evaluation literature:

$$ATT = E[(Y_{1i}|D = 1) - (Y_{0i}|D = 0)]$$

where Y_{1i} denotes the outcome (in our case, the wage) of individual i conditional on participation in the program, while Y_{0i} indicates the alternative outcome for the same individual if he hadn't participated into the program. The standard problem of program evaluation research is that the theorical ATT is not measurable since the counterfactual outcome $(Y_{0i}|D = 0)$ is not observable.

A standard solution to the problem is building up a counterfactual to approximate $(Y_{0i}|D = 0)$. This can be achieved by matching

individuals who are similar according to a number of observable characteristics, X, affecting both participation and outcomes; one method is matching individuals on the basis of their "propensity score" to particpate into the program, $p(X) = \Pr ob(D = 1|X_i)$, conditional on a number of observables. The aim of the matching approach is to build up a treatment and control group consisting of individuals similar along a high number of observable dimensions, such that their participation into the program can be considered random (Conditional Independence Assumption). As shown by Rosenbaum and Rubin (1983), if the conditional independence assumption holds after conditioning on a set X of observables, then it holds also after conditioning on a function of X, the propensity score $p(X)$.

In the case under study, identifying the appropriate treatment and control groups would be equivalent to ask the following counterfactual questions:

1) which wage would have been earned by a student resident in the South, attending university in the South and then working in the North if he had worked in the South (model I, table 4)?

2) which wage would have been earned by a student resident in the South, attending university in the North and working in the North if he had attended the university in the South (model II, table 4)?

The estimation results are provided in Table 5. The ATTs have been obtained using the Stata program "pscore2" by Becker and Ichino (2002).

Table 5 here

The estimation of the propensity score was based on a set of observables relative to individual ability (high school and university leaving grade, type of high school attended), family background (parental education and profession) and local labour market conditions.

Table 5 is self explanatory, and shows that both effects are sizeable and significant. The first one, around 12%, (model I) is similar, in the interpretation, to the regional wage differential usually estimated in the literature (even if unconditional to university attendance), the second one (model II), between 5% and 8% is the wage effect purely due to attending university in the North rather than in the South. These findings have strong policy implications and seem to play in favour of policies subsidizing or enhancing the mobility of university students from Southern to Northern regions.

5 References

Angrist J. and Krueger A. (1999), Empirical strategies in labour economics, in Ashenfelter and Card (eds.), Handbook of Labour Economics.

Becker, S.O., and Ichino, A. (2002), Estimation of average treatment effects based on the propensity scores, The Stata Journal, Vol. 2, pp.358-377.

Boero, G., McKnight, A., Naylor, R. and Smith J. (2001), Graduates and graduate labour markets in the UK and Italy, CRENOS Working paper, no. 01/11.

Bratti, M, and Mancini, L. (2003). Differences in early occupational earnings of UK male graduates by degree subject: evidence from the 1980-1993 USR, Quaderni di Ricerca, Università Politecnica delle Marche, n.189.

Brewer, D., Eide, E and Ehrenberg (1999), Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings, The Journal of Human Resources, Vol.34, No.1 , Winter, pages 104-123.

Brunello, G. and Checchi, D. (2003), School Quality and Family Background in Italy, IZA Discussion paper no. 705.

Card, D. (1993) Using geographic variation in College Proximity to estimate the return to schooling, NBER working paper.

Card, D. (1999), The causal effect of educlation on earnings, in Ashenfelter and Card (eds.), Handbook of Labour Economics.

Checchi D. (2003), The italian educational system: family background and social stratification, mimeo, presented at the ISAE conference "Monitoring Italy".

D'Ambrosio, C. (2001), Social distance between workers: an application to Italian geographic areas, Economics of Transition, Vol.9, no.2, July.

Davia, M. (2004) The transition out of education and initial steps into the labour market in the European Union, IRISS working paper series, no. 2004-01.

Denny, K. and Harmon, C. (2000), The impact of education and training on the labour market experiences of young adults, IFS working paper no. 00/08.

Dolton, P.J., and Makepeace, G.H., (1990), The earnings of economics graduates, The Economic Journal, 100 (March): pp. 237-250.

Dolton, P.J., and Makepeace, G.H., (1986), Sample selection and male-female earnings differentials in the graduate labour market, Oxford Economic Papers.

Flabbi, L. (2001), Secondary schooling decisions in Italy, mimeo, NYU.

Hilmer, M.J. (2001), A comparison of alternative specifications of the college attendance with an extension to two-stage selectivity model, Economics of Education Review, 20, p. 263-278.

Lee, L.F. (1983). Generalized econometric models with selectivity, Econometrica, Vol. 51, No. 2 (March): pp. 507-512.

McKnight, A., Naylor, R., and Smith, J. (2000). Graduate employability: policy and performance in higher education in the UK, The Economic Journal, 110 (June): pp.382-411

Naylor, R., and Smith, J. (2001). Determinants of degree performance in UK universities: a statistical analysis of the 1993 cohort, Oxford Bullettin of Economics and Statistics, 63, 1: pp. 29-60

Rosenbaum, P.R., and Rubin, D.B., (1983). The central role of the propensity score in observational studies for causal effects, Biometrika, 70, 1: pp. 41-55.

Table 1. Sample statistics.

	Females	Males	Total
Degree subject			
math-scientific	535	546	1,081
chemistry-pharmacy	560	471	1,031
geo-biology	681	463	1,144
medical	635	624	1,259
engineering	595	1,775	2,370
architecture	564	575	1,139
agrarian	383	402	785
economics-statistics	1,563	1,655	3,218
political-science	842	674	1,516
law	1,361	971	2,332
humanities	1,349	573	1,922
linguistic	992	263	1,255
teaching	631	245	876
psychology	596	320	916
Total	11,287	9,557	20,844
Economic activity			
employed	8,003	7,514	15,517
unemployed	1,343	590	1,933
out of the labour force	1,812	1,314	3,126
Professional groups			
intellectual professions	3,465	3,712	7,177
technical specialized professions	3,315	2,977	6,292
administrative professions	873	417	1,290
profession with no specialization	336	387	723
Residence during university			
north	5,502	4,776	10,278
centre	2,691	2,318	5,009
south	3,094	2,463	5,557
Residence during work			
north	4,392	4,128	8,520
centre	1,665	1,564	3,229
south	1,520	1,574	3,094
Other individual variables			
age	30.1	31.2	
university leaving grade	104.3	101.7	
high school leaving grade	49.5	48.1	
Average wage by degree subject			
math-scientific	1118.4	1297.7	
chemistry-pharmacy	1108.7	1271.5	
geo-biology	978.5	1145.9	
medical	1421.0	1673.2	
engineering	1268.6	1363.2	
architecture	1032.4	1193.6	
agrarian	1055.5	1172.2	
economics-statistics	1121.3	1318.6	
political-science	1064.9	1328.3	
law	948.7	1089.8	
humanities	931.2	1066.9	
linguistic	959.4	1050.7	
teaching	948.8	1177.1	
psychology	915.7	1123.5	

Table 2. Geographic mobility during university and work.

	Region of residence before university		
Residence during university	North	Centre	South
North	97.3	5.4	7.3
Centre	2.3	92.9	12.4
South	0.4*	1.7	80.3
Total	100	100	100

	Region of residence during university		
Residence during work	North	Centre	South
North	95.9	11.0	10.8
Centre	2.5	80.6	6.9
South	1.6	8.4	82.3
Total	100	100	100

	Region of residence before university		
Residence during work	North	Centre	South
North	98.3	7.8	15.4
Centre	1.3	90.6	10.5
South	0.5*	1.6	74.1
Total	100	100	100

Note: * indicates a number of sample observations below 50.

Table 3. Geographic mobility: transition matrices

	Region of residence before university											
	North				Centre				South			
Residence during university	**Residence during work**				**Residence during work**				**Residence during work**			
	North	Centre	South	Total	North	Centre	South	Total	North	Centre	South	Total
North	98.7	1.0	0.3*	100								
Centre	84.9	13.8*	1.2*	100								
South	-	-	-	-								
North					35.6	63.7	0.6*	100				
Centre					6.2	92.6	1.2*	100				
South					-	-	-	-				
North									62.2	8.2*	29.5	100
Centre									12.8	38.4	48.8	100
South									10.7	6.1	83.2	100

Note: * indicates a number of sample observations below 50.

Table 4. Identification of the treatment and control groups.

Model I: identifying the "Job place effect".

Region of residence			
Before university	**During University**	**During Work**	
South	South	North	(treatment)
South	South	South	(control)

Model II: identifying the "University Location effect".

Region of residence			
Before university	**During University**	**During Work**	
South	North	North	(treatment)
South	South	North	(control)

Table 5. Propensity score matching estimation results: the average effect of the treatment on the treated.

"Job place effect":

Model	Matching Estimator	Treatment (obs.)	Control (obs.)	ATT	% average wage (control)	s.e. (boots.)	t
I	Nearest neighbor	305	284	127.9	12.4	41.00	3.12

"University Location effect":

Model	Matching Estimator	Treatment (obs.)	Control (obs.)	ATT	% average wage (control)	s.e. (boots.)	t
II	Nearest neighbor	196	151	97.28	8.2	49.06	1.98
II	Radius	196	305	61.63	5.2	42.27	1.46
II	Kernel	196	305	103.56	8.7	38.21	2.71