

Contextual factors of the external effectiveness of the university education: a multilevel approach

Matilde Bini¹, Leonardo Grilli², Carla Rampichini²

Abstract This paper aims to measure the external effectiveness of the course programs of Italian universities taking into account both the characteristics of graduates and some context factors that differently affect the Italian regional labour markets. The analysis is performed via a multilevel logistic model on a dataset coming from a survey on job opportunities of the Italian graduates in 2004, conducted by Istat in 2007.

Key words: contextual effects, effectiveness, multilevel data analysis.

1 Introduction

The increasing unemployment rates of young graduates recently observed in many European countries revives interest in studying the transition from university to work. An important aspect of this phenomenon is verifying if the degree of education acquired from university is adequate to the needs of the labour market. Thus, it is necessary to study the effectiveness of the university educational process with respect to the labour market outcomes of graduates, also named external effectiveness. This concept can be measured through several indicators, like the success in getting a job, or a job in conformity with the education supplied by the course program attended by graduate (Bini, 1999), or a job with adequate salary or income, or the time to get the first job (Biggeri *et al.*, 2001). This study deals with having a job about three years

¹ European University of Rome, Via Aldobrandeschi 190 – 00163 Roma; mbini@unier.it

² Dipartimento di Statistica “G. Parenti”, Viale Morgagni, 59 – 50134 Firenze

Acknowledgements: we thank Istat for allowing the use of the data on Italian graduates.

after graduation. Any analysis on the effectiveness of the university educational process should take into account the factors out of the control of the institution that affect the effectiveness indicators: features of the graduates, features of the universities (like for example indices of financial resources, number of students per course program), and external context factors represented by socio-economic features of the region where the university is located and where most of its graduates are going to search for a job. Some studies dealing with this issue in Italy, like the surveys on job opportunities conducted by AlmaLaurea in 2008 (Cammelli, 2009) and by the Italian National Statistical Institute (Istat) in 2007, as well as the analysis carried out by Bini in 1999, revealed that results about Northern regions are different from those about Southern regions. This means that a problem such as getting the first job is due not only to the differences among individuals and among the different capabilities to educate individuals that universities have, but also to the economics and social differences among regions where universities are located. Therefore, a fair comparative effectiveness analysis should include observable variables or suitable proxies measuring the relevant social and economic features of territories. This paper aims at measuring the external effectiveness of the course programs of all universities considering both the characteristics of individuals and context factors that differently affect the regional job markets, in order to make '*coeteris paribus*' comparisons. This kind of analysis could help the government and policy makers to implement strategies for improving the tertiary education and allocate resources. Our analysis exploits the survey on the on job opportunities of the Italian graduates in 2004, conducted by Istat in 2007 (Istat, 2008). Several context indicators can be used in the analysis: *i*) macroeconomic measures, like the Gross Domestic Product per inhabitant, or the productivity of labour; *ii*) job market measures, like the (youth) unemployment rate, or the quota of irregular labour; *iii*) measures of production structures, like the number of firms per inhabitant and the average number of employees per firm; *iv*) innovation and technology measures, like the quota of innovative firms; *v*) measures of the degree of culture, like the quota of family expenses for cultural entertainments; *vi*) quality of life measures, like the poverty rate. Since the mentioned context indicators are quite stable in the years from 2004 to 2007, we simply used the values of the indicators observed in 2007 (Barbieri *et al.* 2008). The analysis is based on a multilevel logistic model that take into account the hierarchical two level structure of the data: each graduate is nested in a cluster on the basis of the disciplinary group of the course program. These groups are defined differently within each university (for example, the Scientific group in Florence is a cluster distinct from the Scientific group in Bologna). These clusters will be called group/university later on.

2 Data

The Istat data set used in the analysis is made up by two distinct stratified samples for males and females, where the strata of each sample were defined as intersections between course programs and universities. The number of interviews is 47300, which is about 18% of the universe of the 260070 graduates of year 2004. Note that the important reform of degree programs dictated by the decree law DM 509/99 started in a.y. 2001/2002, so the sample includes 26570 pre-reform graduates and 20730 post-reform graduates. This peculiarity suggests to accomplish two separate analyses. We focus on post-reform graduates because they come from the new degree programmes that should have affected the ‘performance’ of universities in term of capacity to prepare young people to needs of job market. The available graduate covariates, measured at individual level, include demographic variables (age, sex, marital status, place of residence during the studies, military service, social background), university career variables (kind of the high school attended, occupational condition during the studies, course program changes, kind of degree, final mark), after degree information (involvement in other studies or training jobs after the degree).

Moreover, we considered job market variables coming from external data sources (Barbieri *et al.*, 2008) and measured at regional level. In this preliminary analysis, factors pertaining internal context (features of the degree programs or universities) measured at the group/university level, have not been included yet.

Since we wish to estimate the probability to get the first job, the sample of 47300 records has been reduced to 32228 by eliminating all graduates who, at the date of the interview, did have the same job before their degrees, and who were unemployed but at the same time were not interested in searching for a job. Moreover, the dataset includes all the Italian universities, but the data pertaining to the universities with less than 750 graduates are collapsed, so that we exclude them in our multilevel analysis.

3 The two-level logistic model

Multilevel modelling is a methodology suitable for effectiveness evaluation (Grilli and Rampichini, 2009). In our application, the data have a hierarchical structure with two levels, represented by graduates (level-one units) and by groups of course programs combined with universities (group/university: level-two units). The observed response y_{ij} is a binary indicator which is equal to 1 if the graduate is employed at the date of the interview. A two-level

logistic model for graduate i of group/university j , with binomial variation at level 1, can be defined as follows (Goldstein, 1995):

$$\text{logit}(\pi_{ij}) = \alpha + \sum_h \beta_h x_{hij} + \sum_l \gamma_l z_{lj} + u_j \quad u_j \sim N(0, \sigma_u^2)$$

where π_{ij} is the probability to get job, x_{hij} is the h -th covariate of graduate i of group/university j , z_{lj} is the l -th macro-economic covariate of group/university j , and u_j is the random effect at second level which is intended to collect the unobserved factors at the group/university level affecting the outcome of the graduates. The model includes several individual-level covariates selected from the Istat questionnaire and some macro-economic variables measured at regional level, built from external data, namely the Gross Domestic Product per inhabitant (*gdp_05*), the unemployment rate (*unempl_05*) and the number of firms per inhabitant (*firms_inhab*). These are the most typical indices representing the degree of economic and job market development of a territory. The values of the contextual variables refer to the region where the university is located, so we implicitly assume that the graduates search for job in the region where they got their degrees.

4 Main results

The results from the multilevel logistic model are shown in Tables 1 and 2. Only few individual-level covariates are significant: age of graduates (*age2*, in class: <23, 23-24, 25-29, >29), graduation within institutional time (*graduate_in_time*, 1 if yes), occupational condition during the studies (*working_student*, 1 if at work during the studies), attended courses after getting the degree (*after_degree*, e.g. master or PhD, 1 if yes), educational level of father (*graduate_father*). For the last variable, also the cluster mean (*pctg_graduate_father*) was inserted, showing the education level of the father has a significant contextual effect, while the individual level coefficient is not significant. As regards the macro-economic variables measured in the 2005 year at Regional level, only the unemployment rate (*unempl_05*) is significant. The significant residual variance at level 2 indicates the presence of further unobserved factors at group/university level.

Table 1: Estimates of two-level logistic model: Model A (without external context variables)

	q2_1	Coef.	Std. Err.	P> z
	+			
age2_2		.3238295	.1220217	0.008
age2_3		1.030602	.1650102	0.000
age2_4		.5664397	.2385697	0.018
graduate_in_time		.4763298	.1194794	0.000
working_student		.2533406	.0943595	0.007
graduate_father		-.1923589	.1107814	0.082
pctg_graduate_father		-.0176543	.004933	0.000
after_degree		-.6314676	.0794733	0.000
_cons		1.942942	.3150035	0.000
sigma_u		1.171709	.0875581	

Table 2: Estimates of two-level logistic model: Model B (with external context variables)

	q2_1	Coef.	Std. Err.	P> z
	+			
level 1				
age2_2		.3171903	.1206581	0.009
age2_3		1.080437	.163419	0.000
age2_4		.6364174	.2353766	0.007
graduate_in_time		.4914841	.1189942	0.000
working_student		.2023089	.094447	0.032
graduate_father		-.1924619	.1106952	0.082
after_degree		-.6274822	.0792874	0.000
_cons		2.838565	.3464269	0.000
level 2				
pctg_graduate_father		-.0201524	.0047053	0.000
unempl_05		-.1372708	.0220267	0.000
sigma_u		1.0432	.0839319	

Comparing Model A and Model B, we can see that the unexplained level 2 variance reduces of about 21% after the introduction of the external contextual variable *unempl_05*.

Model B can be used to predict the probability of being employed for a baseline graduate (22 years old at graduation, graduated within institutional time, with a non graduated father, who never attended courses after graduation, and was not working during the studies) in a hypothetical group/university with *pctg_graduate_father*=0. Since the random effect u_j is normally distributed with an estimated standard deviation of about 1.04, we define a group/university as very good if $u_j = 2*1.04$, good if $u_j = 1.04$, medium if $u_j = 0$, bad if $u_j = -1.04$, and very bad if $u_j = -2*1.04$. Figure 1 shows the predicted probabilities for a baseline graduate in Lombardy, Tuscany, Campania and Sicily regions using the corresponding values of the contextual variables, and plugging in the above values of the random effect.

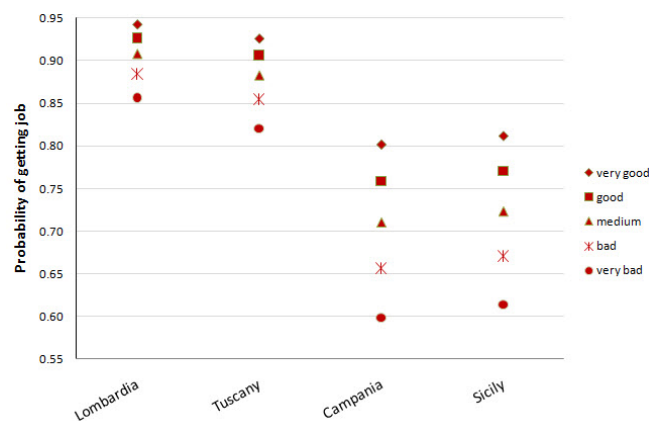
The results show that a bad group/university located in a region with a favorable economic context is better, as for occupational chances, than a good group/university located in a region with an unfavorable economic context. This implies that a fair comparison should take into account the factors that characterize regions and affect external outcomes of the universities.

We then computed level 2 residuals per geographical macro area (North, Centre, South) and group of disciplines (9 groups: Humanities, Economics and

Statistics, Political and Social Sciences, Science, Law, Engineering, Architecture, Medical sciences and Physical education). Note that we considered only first level degree programs (3 years), so excluding the typical degree programs in Law, Architecture and Medical sciences, lasting 5 or 6 years. The residual analysis showed that degree programs in Humanities, Political and Social sciences, Law, Science and Architecture, yield lower chances of getting a job independently from the geographical area; while only Medical sciences (mainly nursery) yield higher chances in all the geographical areas. In general the residuals vary across geographical areas.

This means that labor market dynamics could completely or partially affect some disciplines, so that comparisons among universities performances where contextual factors are taken into account, should be done with caution, as well as it should be better identify variables that exactly measure the peculiarities of each region (or geographic macro area).

Fig. 1 Predicted probabilities of getting job for graduates



References

1. A. Cammelli (a cura di): AlmaLaurea. XI Rapporto sulla condizione occupazionale dei laureati. Occupazione e occupabilità dei laureati a dieci anni dalla Dichiarazione di Bologna. (2009)
2. Biggeri L., Bini M., Grilli L.: The transition from university to work. A multilevel approach to the analysis of the time to obtain the first job, *Journal of the Royal Statistical Society, series A*, **164**, 293-305 (2001).
3. Bini M.: Valutazione della Efficacia dell'Istruzione Universitaria rispetto al Mercato del Lavoro. In: Research Report n.3/99, Osservatorio per la Valutazione del Sistema Universitario, MIUR, Roma, 1-100 (1999).

4. Goldstein, H.: Multilevel Statistical Models, Edward Arnold, London (1995).
5. Grilli L., Rampichini C.: Multilevel models for the evaluation of educational institutions: a review. In: Bini M, Monari P, Piccolo D, Salmaso L (Eds) Statistical methods for the evaluation of educational services and quality of products. Physica-Verlag, 61-80 (2009).
6. ISTAT: Inserimento professionale dei laureati. Indagine 2007. ISTAT, Roma (2008).
7. Barbieri A G., Cruciani S., Ferrara A. (a cura di): 100 Statistiche per il Paese. Indicatori per conoscere e valutare. ISTAT, Stampa CSR (2008).