The literacy divide: territorial differences in the Italian education system

Claudio Quintano, Rosalia Castellano, Sergio Longobardi
University of Naples "Parthenope" Via Medina 40 , Naples,
e-mail: claudio.quintano@uniparthenope.it , lia.castellano@uniparthenope.it,
sergio.longobardi@uniparthenope.it

Abstract: The aim of this work is to analyze the differences in student achievement across Italian macro regions. We provide a multilevel analysis of data from the last PISA (Programme for International Student Assessment) survey edition (2006) and investigate how several factors may influence the results of the educational process. In addition to the classic determinants of student achievement (e.g. family socio-economic background) the goal of this paper is to investigate the extent to which the test-taking motivation may contribute to influence the results from assessment test and to explain partially the territorial differences.

Keywords: statistical methods for evaluation, multilevel model, regional inequality

1. The Italian literacy divide

The PISA (Programme for International Student Assessment) survey, which takes place every three years, is carried out by the Organization of Economic Cooperation and Development (OECD) and it has been designed to assess the skills and knowledge of 15-year-old students in three main areas –reading, mathematics and science-. Fifty-seven countries participated in the last edition of PISA (2006), including all 30 OECD countries. In Italy, approximately 22,000 15-year-olds from about 800 schools participated. The Italian students in PISA 2006 reach an average test score of 462 points in mathematics, 469 in reading and 475 in science, being under the OECD average of 500. The gap between Italian students and top performing countries like Korea and Finland is extremely high and Italy performs significantly worse than all OECD countries, excepting the Republic of Slovak, Turkey, Spain and Greece. The very poor performance of Italian students is due to significant territorial differences within the country. Indeed, fifteen year-old students in the Italian Southern regions performed very low in each assessment area which contributed to Italy's standing in international comparisons. For each PISA cycle (2000, 2003, and 2006) the average score differs strongly among the Northern and the Southern regions and these marked differences originate a wide North-South divide which is called *literacy divide.* Many studies (Marks, 2006; Korupp *et al*., 2002) emphasize the role of socio-economic background for determining learning outcomes and explaining the territorial differences[1] while this paper aims to asses how much of the Italian literacy divide is attributable to differences in the test-taking motivation. Consequently, we propose a two stage approach: firstly, a latent variable which expresses the student engagement and effort is developed,

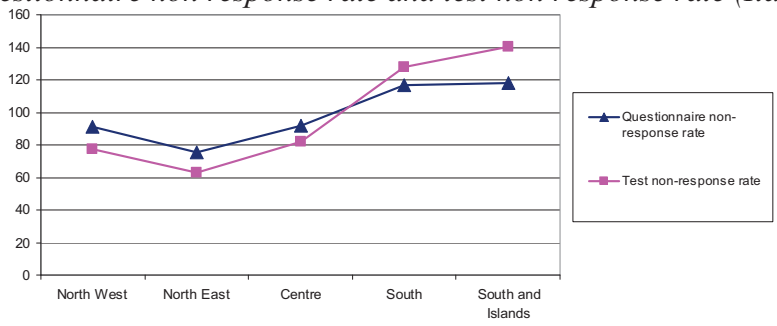
[1] With reference to Italian context this topic is also discussed in Quintano, Castellano, Longobardi (2009).

secondly, a multilevel regression is performed to analyze the effect of this factor on student performance after controlling for school and student variables.

2. The students' effort

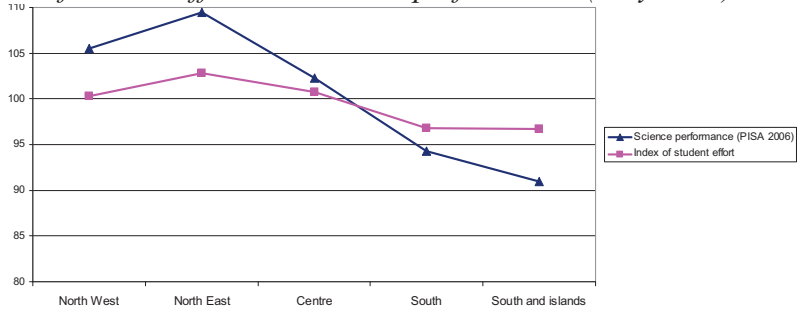
The PISA test is considered as a *low stake test* since the students perceive an absence of personal consequences associated with their test performance. Without an adequate effort, test performance is likely to suffer, resulting in the examinee's test score underestimating his or her actual level of proficiency (Wise and De Mars, 2005). Three variables are considered to express the students'test motivation: the "*Test non-response rate*" computed on the basis of the number of missing or invalid answers in the PISA cognitive test. This ratio could be influenced by the competence of the student since if a student does not know the correct answer he could skip the question (missing response) for this reason as second variable we consider the "*Questionnaire non-response rate*" computed on the basis of the number of missing or invalid answers in the PISA student questionnaire which are not related with the proficiency level of the student. The test non-response rate shows a territorial trend similar to questionnaire non-response rate (Fig.1) confirming that both these ratios could be useful to express the response behaviour of the students.

Figure 1: *Questionnaire non response rate and test non response rate (Italy=100)*



The third variable is the students'self-report effort in the PISA questionnaire measured on a 10-point scale. These three variables are summarized, by factor analysis, into one latent variable called "Index of student effort" which expresses the effort of the student in the test. The chart (fig.2) highlights the variation of the index of student effort in correspondence of the average performance at macro region level. The correlation between this index and the science performance is equal to 0,553 at national level.

Figure 2: *Index of student effort and science performance (Italy=100)*



These empirical evidences lead us to suppose that the southern students are unawareness of the PISA test importance; consequently it seems that their weak motivation plays an important role to determine low test scores.

3. The multilevel model

In order to identify the possible determinants of the Italian student achievement a multilevel regression model with random intercept is applied[2]. The choice of the multilevel approach is suggested by the hierarchical structure of the PISA data where students (level-one units) are nested in schools (level-two units). The two-level random intercept regression model for the i_{th} student in the j_{th} school can be written as follows:

$$Y_{ij} = \gamma_0 + \sum_{k=1}^{m} \beta_k x_{kij} + \sum_{t=1}^{s} \beta_t z_{tj} + \varepsilon_{ij} + U_{oj} \qquad (1)$$

Where x_k are m variables at student level and z_t are the s variables at school level while ε_{ij} and U_{oj} denote the error components respectively at students and school level, these components are supposed to be normally distributed and uncorrelated:

$$\varepsilon_{ij} \sim \text{IID-N}(0, \sigma^2) \qquad U_{0j} \sim \text{IID-N}(0, \tau^2) \qquad cov(U_{0j}, \varepsilon_{ij}) = 0 \qquad (2)$$

The model requires a preliminary estimate of the empty model (model with no independent variables) to split the total variation on the dependent variable into within and between variance, then a *block entry* approach is adopted (Choen and Choen, 1983) which consists to the gradual addition of the first and second level covariates.

4. Main results and concluding remarks

The proposed multilevel analysis has required seven models (table 1) to compare the impacts of individual and school characteristics including in the last model the index of student effort. The dependent variable is the student science achievement measured by five scaled scores called *plausible values*[3]. The set of independent variables is composed by 6 student-level and 8 school-level characteristics derived from the PISA questionnaires. After the empty model, the second model included variables describing the gender of students, the immigration status, the level of home educational resources and the hours per week spent doing homework. It shows a gender gap of 12 points in favour of males and a gap of 42 points in favour of non immigrants. Once the macro area dummies are introduced (fifth model), a comparison with the previous model shows that the socioeconomic gap is narrowed down from 91 points to 34 points while the gap associated with the private schools increased from 49 points to 56 points, also the differences correlated with the school type are increased, for example the gap of lower secondary students raises 120 points. The last multilevel model brings in the index of student effort as control variable. This variable shows a high and significant impact on student performance and this factor involves the reduction of the macro area coefficients. Furthermore, the index of student effort allows to explain a larger amount of variance, indeed after controlling for students motivation the accounted total variance

[2] See Raudenbush and Bryk (2002), *inter alia*, for a relevant discussion on multilevel models.

[3] The multilevel analyses proposed in this paper are developed by the Hierachical Linear Model (HLM) software (Raudenbush, Bryk, Cheong, & Congdon, 2000) in order to handle plausible values as the dependent variable. The three continuous variables at student level are centred on the school mean while the four continuous variables at school level are centred on the grand mean.

among students (compared with the sixth model) increases from 4% to 20% while the accounted total variance among schools increases from 80% to 87%. Finally, our findings highlight that the North-South divide has been overestimated by the PISA test since the score differences are also influenced by the lower effort and engagement of the southern students. moreover, the multilevel analysis has confirmed the role of the socio-economic context to influence the student achievement.

Table 1: *Estimation of multilevel regression coefficients*

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Intercept	460.09***	466.35***	468.13***	495.24***	511.97***	514.31***	508.35***
Student level variables							
Gender (reference=male)		-11.64***	-11.9***	-12.20***	-12.34***	-12.25***	-12.12***
Immigrate (reference=native)		-42.44***	-41.78***	-39.98***	-42.45***	-42.47***	-39.46***
Index of Home educational resources		6.72***	6.73***	6.79***	6.73***	6.74***	4.50***
Index of self-confidence in ICT		4.71***	4.71***	4.69***	4.69***	4.69***	3.79***
Hours per week spent on homework (reference=0-2)							
0		-10.52***	-11.21***	-10.86***	-10.95***	-10.88***	-3.56
02-04		7.19***	6.77***	6.47***	6.62***	6.55***	1.36
>4		9.93***	9.52***	9.12***	9.79***	9.68***	3.14
Index of student effort							32.45***
School level variables							
Scholastic context							
Index of economic, social and cultural status -ESCS- (school average)			91.16***	60.95***	34.07***	36.43***	33.62***
Parents' pressure academic standards (reference=low pressure)			-0.86***	-1.77	1.82	4.48	2.77
Private school (reference=public school)				-49.17***	-56.17***	-50.58***	-45.00***
Study programme (reference= schools specializing in classical studies or in science education							
Technical institutes				-12.76*	-30.71***	-33.47***	-30.62***
Professional institutes				-49.48***	-72.43***	-70.64***	-55.72***
Vocational training				-26.64	-87.01***	-89.63	-68.04***
Lower secondary school				-99.28***	---	---	-86.11***
Macro area (reference=Center)							
North East					42.46***	36.49***	30.27***
North West					27.32***	23.15***	22.91***
South					-31.02***	-31.69***	-20.34***
South and islands					-37.78***	-39.14***	-25.65***
Scholastic resources							
Computers with web						0.08***	0.06**
Quality of educational resources						0.31	-0.76
Teacher shortage						7.74***	3.93***
Variance components							
Variance between schools	5,347.98	5,137.21	2,496.45	1,952.86	1,085.71	1,025.56	706.84
Variance within schools	4,674.31	4,462.54	4467.4	4,464.34	4,470.45	4,470.30	3,717.59

Significance level: *** 99%; ** 95%;* 90%

References

Choen J., Choen P. (1983) *Applied multiple regression/correlation analysis for the behavioral sciences*, Hillsdale, New Jersey, Lawrence Erlbaum.

Korupp S.E., Ganzeboom H.B.G., Van Der Lippe T. (2002) Do Mothers Matter? A Comparison of Models of the Influence of Mothers' and Fathers' Educational and Occupational Status on Children's Educational Attainment, *Quality & Quantity*, 36.

Marks G. N. (2006) Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries, *Educational Research* 48,1, pp. 21-40.

Quintano C., Castellano R., Longobardi S. (2009) L'influenza dei fattori socio economici sulle competenze degli studenti italiani. Un'analisi multilevel dei dati PISA 2006, *Rivista di Economia e Statistica del Territorio*, forthcoming.

Raudenbush S.W., Bryk A.S. (2002) *Hierarchical Linear Models: Applications and data analysis methods*, Thousand Oaks, CA: Sage.

Wise S. L., DeMars C. E. (2005) Low examinee effort in low-stakes assessment: Problems and potential solutions, *Educational Assessment*, 10, 1–17.