School Value Added in Italy: evidence from a Primary education level employing different measures.

Adalgiso Amendola [1] ·Alessandra Amendola[2] · Marinella Boccia [3]

July 30, 2021

Draft Version

Abstract

The measure of the Education quality is fundamental to provide information to the public opinion and to give the policymakers the opportunity to implement policies aimed to improve the education supply and the teaching and learning process. To find models that provide a valid measure of the effectiveness of school quality and of Education Value Added (VA) is crucial for the Accountability. The aim of this paper is to compute and evaluate a VA measure for Italian Primary school and compare the performance of different econometric approaches.
Moreover, considering the usefulness of VA measures in evaluate the school efficiency, generating an incentive to cheating or teaching the test and then affecting its results, the role of monitoring is also investigated.
The study will use the INVALSI standardized tests in Italian for the cohort of students at grade II (Primary) in 2011-12; 2012-13 and the cohort of students at grade V-(Primary) in 2014-15 2015-16.
The reached results make in evidence that higher test scores are driven by the test scores in the previous period and the parents' educational level also play a role. According to the expectations, a significant negative sign coefficient is estimated for the presence of monitoring. The School Value Added distribution have also derived for different areas.

Keywords: Value added, Accountability, Primary education, Heterogeneity of value-added distribution

[1] DiSES-University of Salerno; Via Giovanni Paolo II, 132- Fisciano, 84084-Salerno -Italy; adamendola@unisa.it
[2] DiSES-University of Salerno; Via Giovanni Paolo II, 132- Fisciano, 84084-Salerno –Italy; alamendola@unisa.it
[3] DiSES-University of Salerno; Via Giovanni Paolo II, 132- Fisciano, 84084-Salerno -Italy; mboccia@unisa.it

1. Introduction and literature

The measure of the Education quality is fundamental not only in order to provide information to the public opinion, but also to give the policymakers the opportunity to implement policies aimed to improve the education supply and the teaching and learning process.

To find models that provide a valid measure of the effectiveness of school quality and of education Value Added (VA) is crucial for the Accountability. In Italy, during the last years, the Accountability has become of special interest. The first reason for that is related to the schools' possibility to choose their internal organization, such as the teaching and learning process; the second reason meets the importance of the school evaluation in the sense that it can be useful to help both parents, in choosing better schools for their children, and the policymakers, in adopting the relative policies aimed to improve, if necessary, the quality of education system and allocate the resources in an efficient and effective way. Given that, the measure of the school effectiveness has become a fundamental question (Jacob, 2002; Figlio and Rouse, 2005; Figlio and Lucas, 2004; Figlio and Kenny, 2009; Holbein and Ladd, 2017) and the measure of the value added a crucial point.

Value-added models of teacher quality were first introduced in the seventies by Hanushek (1971) and Murnane (1975). More recent papers include Rockoff (2004), Rivkin, Hanushek, and Kain (2005), Aaronson, Barrow, and Sander (2007), and Kane and Staiger (2008).

The literature provides several measures of value added and some papers focused on the context of teacher value-added: among them Chetty et al. (2014) (a) and (b) found that VA models, which control for a student's test scores in the previous period, give an unbiased forecast of teachers impacts on student achievement. The discussion on the methodology is also another issue: Everson (2017) in a recent study deal with the appropriate modelling choices, De Simone and Gavosto (2013) managed to estimate school effects on cognitive gains. Based on students' longitudinal test scores, for 72 schools involved in an experiment by the Italian Ministry of Education, they used several econometric approaches, that give robust results.

In 2017 Agasisti and Mynaia, in line with Riehl et al. (2016), estimated school performance by adopting different models that include different combinations of methods and covariate selection to adjust for compositional differences across schools.

The aim of this paper is to derive an unbiased measure of Value Added in Italy and highlight the difference that can characterize the Italian geographic areas. Moreover, because the use of VA measures is helpful in order to evaluate the school quality and this could encourage reaction such as teaching or cheating during the tests (see Bertoni et al.(2013) and Angrist et al. (2017) for a discussion), changing the signal in VA measures, the analysis will investigate the role of cheating across the country. The empirical study is based on INVALSI standardized tests in Italian[4] for primary school and in particular it will follow two cohorts of students at grade II (primary) in 2011-12; 2012-13 and the two cohorts of students at grade V-(Primary) in 204-15; 2015-16. This allow to compute panel data at school level.

The main research interest is to provide an answer to the following hypothesis: H1: What are the determinants of Test score in the primary school? ; H2: are the results robust to the different models employed?; H3: How the measure of valued added differs among the model? And is there difference between the geographic macro areas considered?

The preliminary reached results make in evidence that higher Italian test scores are driven by the test scores in the previous period and are related to the parents' educational level. Remarkable, the estimation for the coefficient indicating the presence of monitoring is significant and with negative sign negative in according to the expectations.

From what conger the geographical analysis the School Value Added distributions are derived for selected representative areas According to these the North shows a higher density on the right and side of distribution, differently from that the South that exhibits lower density in the VA distribution.

The rest of the paper is organized as follows, after an introduction and a brief review of literature in the section 2, we will explain the data object allowing for a preliminary analysis. Section 3 provides details about the methodology. In section 4, results and conclusions related to each model are reported highlights the main finding related to the determinants of test and the characteristics of different geographic areas. Finally, Section 5 concludes.

[4] In this preliminary version the estimation will consider only the italian test score. The idea is to compute the same estimation for math and discuss about the differences.

2. Institutional Details and Data

In 2008 in Italian primary school were introduced the INVALSI standardized tests in Italian and Math. The aim of these tests was to evaluate school performance and inform the authority about the effectiveness of the school system. These tests also offer schools a consistent reference to their assets and not, using a value added approach. Of course all the agents involved are interested to perform well. First of all because they might be used to evaluate teachers and schools and secondly because results affect the school's reputation even if the results of tests score are not made public by INVALSI. From 2008 the tests have been dispensed every year. In this paper, we start with the 2010 for several reasons. First, this wave was the first to administer the test and collect data for all the students of Italian primary school in their second and fifth grade. Furthermore, in the sampled classes the test was done in the presence of an external examiner. The monitoring that had to be present in the class during the test and supervisor its correct accomplishment and report the answers on the dedicated answer sheets and transfer them to INVALSI. In the other classes, the test was administered by teachers of the school (but not of the class and not in the subject tested), and reporting was done jointly with the teacher of the class (see also Bertoni et al. 2013 for a discussion).

The aim of this paper is to use the INVALSI data standardized tests in Italian and Math for primary school and in particular in this version the analysis is computed, at school level, considering the Italian test score and following two cohorts of students at grade II (primary) in 2011-12; 2012-13 and the two cohorts of students at grade V-(Primary) in 2014-15; 2015-16 for a total of 6,195 Schools. The table 1 below shows the descriptive statistic related, realized on the average value of all the variables at school level, while Table 2 reports the correlation between the variables of interest. At same time Figure 1 shows the Italian test score distribution (total and by areas), both at V (panel a, b) and at II level (panel c, d). In both graphs we observe a higher density in the upper tail of distribution especially in Nord east and Nord west, differently from that in the South where the tails density is lower.

Table 1: Descriptive Statistics

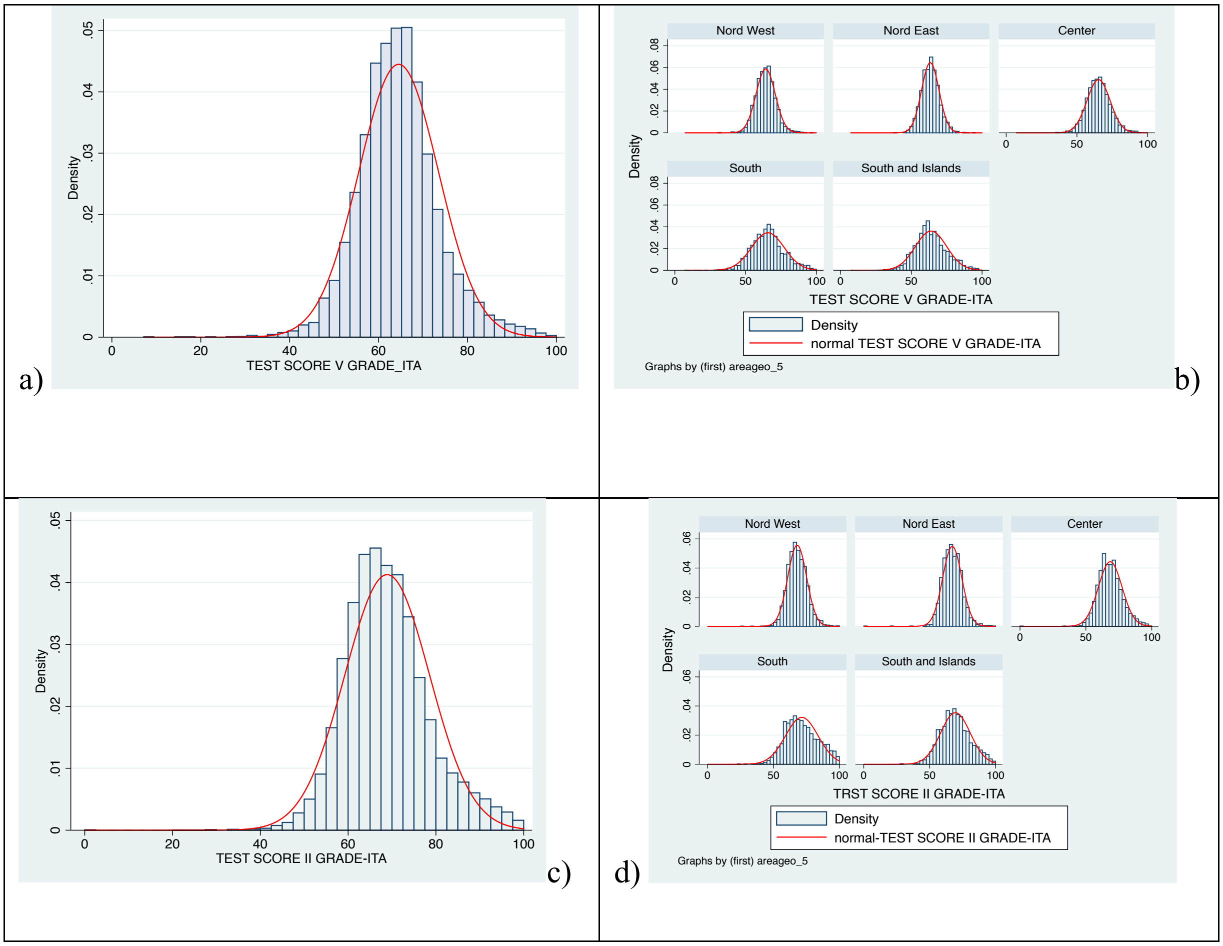
Variable	Obs	Mean	Std. Dev	Min	Max
Test Score V	11,302	64.5519	8.898538	15.71429	100
Test Score II	11,302	68.8519	9.656394	0	100
Monitoring	11,302	.0377014	.1274353	0	.9882353
Number of students	11,302	17.74711	4.286696	1	33
Females	11,302	.4878503	.093613	0	.9
Father Higher Education	11,302	.1353166	.1412262	0	.9565217
Father higher Occupation	11,302	.0298037	.0505662	0	.6206896
Mother higher Education	11,302	.1647053	.1461557	0	.9090909
Mother higher Occupation	11,302	.0106626	.024974	0	.3157895
Advanced Student	11,302	.0240111	.0367757	0	.8
1st generation immigrant student	11,302	.0277793	.0377596	0	.7142857
2 st generation immigrant student	11,302	.0621923	.0722812	0	.9333333

Table 2:Correlation Matrix

Variables	Test Score V	Test Score II	Monitoring	Number of students	Females	Father Higher Education	Mother higher Education	Mother higher Education	Mother higher Occupation	Advanced Student	1st gen immigrant student	2 st gen immigrant student
Test Score V	1.0000											
Test Score II	0.2926	1.0000										
Monitoring	-0.0939	-0.0548	1.0000									
Number of students	0.0844	-0.0163	-0.0067	1.0000								
Females	0.0278	-0.0023	0.0051	0.0755	1.0000							
Father Higher Education	0.1961	0.1314	-0.0571	0.2639	0.0207	1.0000						
Father higher Occupation	0.1224	0.0714	-0.0469	0.1971	-0.0064	0.6406	1.0000					
Mother higher Education	0.2024	0.1119	-0.0477	0.2630	0.0200	0.8903	0.5898	1.0000				
Mother higher Occupation	0.1180	0.0647	-0.0290	0.1588	-0.0055	0.5190	0.6008	0.5206	1.0000			
Advanced Student	-0.1029	-0.0453	0.0075	-0.1205	-0.0204	-0.0827	-0.0684	-0.0908	-0.0593	1.0000		
1st gen immigrant student	-0.1424	-0.1215	0.0297	-0.0929	-0.0025	-0.1332	-0.0746	-0.1227	-0.0820	0.3712	1.0000	
2 st gen immigrant student	-0.1254	-0.1859	0.0238	0.0199	0.0269	-0.0874	-0.0441	-0.0781	-0.0532	0.2277	0.4305	1.0000

Figure 1 : Italian Test Score Distribution- V Grade and II Grade - By Total and Macro Area

.



3. Methodology

According to Schagen and Hutchison, (2003) at least three different meanings of the term 'value-added' have become current, namely: 1. measures of 'pure progress', controlling only for each pupil's prior attainment; 2. measures controlling for prior attainment in addition to a range of other pupil- and school-level factors which apparently impinge on pupil progress and are outside the control of the school; 3. measures controlling only for background but not for prior attainment.

In this study we refer to the second definition for the value-added mining and give evidence of the estimated results for different econometrics approaches.

Following part of the literature (i.e. De Simone and Gavosto, 2013) we start from the following linear equation.

$$TestScoreV_{sc}^{Ita} = \alpha_1 TestScoreII_{sc}^{Ita} + \gamma_c + \varepsilon_{sc} \quad (1)$$

The analysis is realized at school level *(s)* on the Italian test score (*Ita*) and considering the cohorts (*c*) as detailed above.

As for the variables of interest we have:

-$TestScoreV_s^{Ita}$, the depended variable computed considering the Italian Test score corrected for cheating (1-Cheating Correction)* Punteggio grezzo for school *(s)* and cohorts (*c*);

-$TestScoreII_{sc}^{ita}$ the Test Score in the previous period*;*

-ε_{sc} is an error term assumed to be *i.i.d* with zero mean and constant variance of σ^2.

Differently from de Simone and Gavosto (2013), we compute the analysis at a school level and we observe a linear relationship between the residuals and the scores, so we don't include any non-linear terms in the regression , such as the square value of test score in the previous period.

In addition, in order to estimate the cognitive progress, we have to considerer the observables characteristics of students and their family and other factors-non managed by schools- that may "touch" the measure of value added. Given that equation (2) is estimated.

$$TestScoreV_{sc}^{Ita} = \alpha_1 TestScoreII_{sc}^{Ita} + \beta_1 StudentsandFamilyCharcteristics_{sc} + \gamma_c + \varepsilon_{sc} \quad (2)$$

In addition to the variables which before we add:

-$StudentsandFamilyCharcteristics_{sc}$ the vector of Student's Characteristics: Sex (percentage of Females) , Citizen (1[st] and 2[st] generation immigrant students) advanced student; Family Characteristics: Farther 's and Mother' s higher occupational and education level. We also control for the presence of monitoring.

The school value added is here computed as the residual terms, the difference between observed achievements and predicted achievements[5]. So the following equation is estimated

$$VA= Ave\ (\varepsilon_{sc}\ )= ave\ (TestScoreV_{sc}^{Ita} - \hat{}\ TestScoreV_{sc}^{Ita}\ ) \qquad (3)$$

In order to take into account some problems of bias that might derive by omitted variable, in equation (3) , schools' level fixed effects (μ_s) are included as follows:

$$TestScoreV_{sc}^{Ita} = \alpha_1 TestScoreII_{sc}^{Ita} + \beta_1 StudentsandFamilyCharcteristics_{sc} + \gamma_c + \mu_s + u_{sc} \quad (4)$$

We decompose the residual from this regression into a school-specific term: μ_C and an idiosyncratic component. Our school performance metric is the coefficient of school fixed effects from equation (4). Finally, we adopt a random effects approach, which on paper is more efficient,[6] , as it has a structure of errors which is apparently closer to the reality of schools. In fact, students are grouped into classes that are in turn nested into schools. Moreover, the random effect approach allows us to capture the sign of localization in that geographic areas.

However, to be consistent, random effect estimates require school effects to be uncorrelated with the explanatory variables at the school level: a condition which is often hard to fulfil.

A typical formulation of such models is:

$$TestScoreV_{sc}^{Ita} = \alpha_1 TestScoreII_{sc}^{Ita} + \beta_1 StudentsandFamilyCharcteristics_{sc} + \gamma_c + \eta_s + \varepsilon_{sc} \quad (5a)$$

$$Where\ \ \eta_s = A + \sigma_s \quad (5b)$$

In this model each school effect consists of two parts:

- -a grand mean (A):
- -an idiosyncratic component σ_s

5 The diffrence between the equation (1) and (2) is that the first one is estimated without controls variarle.

6 The random effect approach allows us also to control for the geographic localization at school level. Fo this reason the dummy for region are added.

Thus, the deviations from the general mean σ_s are taken as estimates of school value- added.

4 Results

The results are reported in Table 3 in which we observe the estimation related to all the four model considered: Model 1 based on equation (1); Model 2 obtained from the estimation of equation (2) with family and schools' characteristics, and Model 3 realized on the basis of equation (3) computed adding the areas' fixed effects and finally Model (4), based on Random effect estimation.

Before discussing the results, it is important to note the stability of the estimated coefficients across different specifications. In particular, the most of coefficients on individual variables do not vary as we move from the model 2 to the fixed effect model 2 :this meaning that omitted variables bias is not an issue in our analysis. In addition, the coefficients and also the significance remain stable also when we adopt a random effect (Model 4) especially if we compare this model with the OLS.

Looking at the results we have, for all the estimates made on, that the test score at V level are positively driven by the test score in the previous period. The presence of monitoring, as we expect, negatively affect the scores. Moreover, a higher percentage of females in the school is a positive determinant such as the father and mother higher educational level that positively affect the score both in the OLS estimation (Model 2) and random effect estimation (Model 4). Considering this last two models, a higher percentage of Advanced Students (*Posticipatario*), negatively affect the results. As to the first generation immigrant students, the more is the presence in the school the lower are the scores. Finally, the second generation of immigrant students has negative effect only in the fixed effect estimation.

However, the most important tool derived is the School Value Added distribution obtained estimating the equation (4) , plotted in the graphs of Figure 2. According to these the North (East, West) and Center plots show a higher density on the right and side of distribution, differently from that of the South and Islands that exhibits lower density in the upper tail of VA distribution. More in detail considering the random effect approach (equation 5) the distribution of value added exhibit an higher density as to Nord regions of Italy (Figure 3).

Table 3: Test Score Estimations-Italian

	MODEL 1	MODEL 2	MODEL 3	MODEL 4
VARIABLES	Test Score V	Test Score V	Test Score V	Test Score V
Test Score II	0.432***	0.407***	0.213***	0.407***
	(0.0113)	(0.0113)	(0.0184)	(0.0113)
Monitoring		-4.765***	-3.501***	-4.765***
		(0.368)	(0.562)	(0.368)
Number of students		0.00467	-0.00429	0.00467
		(0.0234)	(0.0396)	(0.0234)
Females		1.910*	2.182*	1.910*
		(1.044)	(1.316)	(1.044)
Father Higher Education		2.406*	3.096	2.406*
		(1.326)	(1.903)	(1.326)
Father higher Occupation		2.316	3.218	2.316
		(1.811)	(2.693)	(1.811)
Mother higher Education		5.417***	2.897	5.417***
		(1.291)	(1.870)	(1.291)
Mother higher Occupation		4.491	9.071*	4.491
		(3.576)	(4.905)	(3.576)
Advanced Student		-8.999**	-2.672	-8.999**
		(3.640)	(4.490)	(3.640)
1st generation immigrant student		-11.82***	-15.57**	-11.82***
		(4.038)	(6.089)	(4.038)
2 st generation immigrant student		-1.422	-5.403**	-1.422
		(1.272)	(2.333)	(1.272)
Constant	29.64***	30.00***	44.59***	30.00***
	(0.814)	(1.043)	(1.717)	(1.043)
Observations	11,484	11,484	11,484	11,484
Number of codice_scuola	6,507	6,507	6,507	6,507
Method	OLS	OLS	FE	RE

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 2: Value Added (Italian) total and across Areas-Fixed Effects approach

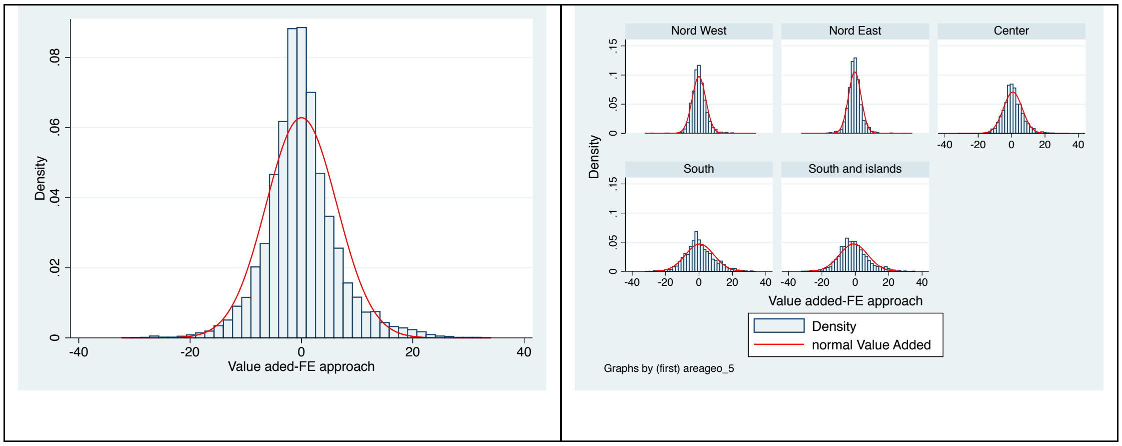
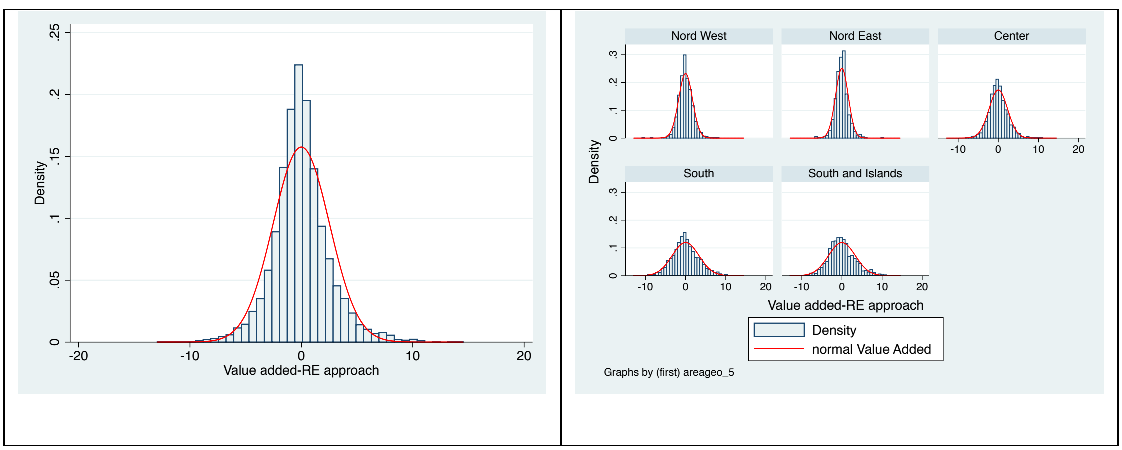


Figure 3: Value Added (Italian) total and across Areas-Random Effects approach



4. Conclusions

The aim of this paper was to estimate the Value Added and in particular derive an unbiased measure of VA in Italy and show the variation between the Italian geographic areas. The study will use the INVALSI standardized tests in Italian and Maths for primary school and in particular it will follow two cohorts of students at grade II (primary) in 2011-12; 2012-13 and the two cohorts of students at grade V-(Primary) in 204-15; 2015-16 (the II primary in 2011-12 will be the V primary in 2014-15 and so on). More in detail we want to answer to the following hypothesis: H1: What are the determinants of Test score in the primary school? ; H2: are the results robust to the different models employed?; H3: How the measure of valued added differs among the models? And is there difference between the geographic macro areas considered?

In order to do that, following De Simone and Gavosto, (2013); Agasisti and Mynaia (2017); Riehl et al. (2016), we adopt different econometric models.

The results, related to Italian test score, are robust to the different models made on, and underline that higher Italian test scores are driven by the test scores in the previous period, such as by the parents' educational level. Remarkable, negative and significant, in according to the expectations, is the coefficient indicating the presence of monitoring and the presence of fist grade immigrant students.

However, the most important tool derived is the School Value Added distribution realized for representative areas. According to these the North shows a higher density on the right and side of distribution, differently from that the South that exhibits lower density in the VA distribution.

References

[1] Agasisti, T. and V. Minaya, (2017). Evluating the Stability of School Performance estimates for school choice: evidence for Italian Primary Schools. SIEP Working paper 733

.

[2] Aaronson, D., Barrow,L. and W Sander. (2007). Teachers and Student Achievement in Chicago Public High Schools. Journal of Labor Economics 25 (1): 95135.

[3] Angrist, J, Battistin E. and D. Vuri (2017). In a Small Moment: Class Size and Moral Hazard in the Italian Mezzogiorno. American Economic Journal: Applied Economics 9 (4) 216-49.

[4] Bertoni, M., G. Brunello, and L. Rocco (2013). When the Cat Is Near, the Mice Wont Play: The Effect of External Examiners in Italian Schools. Journal of Public Economics, 104, 65 77.

[5] Chetty, R, Friedman, J,N. and J. Rockoff. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates: Dataset. American Economic Review. http:// dx.doi.org/10.1257/aer.104.9.2593.

[6] Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. (2014). Measuring the Impacts of Teachers

II: Teacher Value-Added and Student Outcomes in Adulthood.. American Economic Review 104 (9): 263379.

[7] De Simone, G. and A. Gavosto (2013). Patterns of Value-Added Creation in the Transition from Primary to Lower Secondary Education in Italy. PRO- GRAMMA EDUCATION FGA WORKING PAPER N. 47(04/2013).

[8] Everson, K. C. (2017). Value-Added Modeling and Educational Accountability: Are We Answering the Real Questions?. Review of Educational Re- search, 87(1), 35-70.

[9] Figlio, David N, and Kenny, Lawrence W. (2009). Public Sector Performance Measurement and Stakeholder Support. Journal of Public Economics 93(9- 10): 1069-77.

[10] Figlio, David N., and Lucas M. (2004). What's in a Grade? School Report Cards and the Housing Market. American Economic Review 94(3): 591-604.

[11] Figlio, David N., and Rouse, C. (2006). Do Accountability and Voucher Threats Improve Low-Performing Schools?.Journal of Public Economics 90 (1-2): 239-55.

[12] Hanushek, E. (1971). Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data. American Economic Review 61 (2): 28088.

[13] Holbein J. and Ladd H. (2017). Accountability pressure: Regression discon- tinuity estimates of how No Child Left Behind influenced student behavior. Economics of Education Review 58 (2017) 55-67.

[14] Jacob, B. A. (2002). Where the boys arent: Non-cognitive skills, returns to school and the gender gap in higher education. Economics of Education review, 21 (6), 589598.

[15] Jacob, B. A., (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. Journal of Public Eco- nomics, 89(5), pp. 761-796.

[16] Jacob, B. A. and Levitt, S., (2003). Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. Quarterly Journal of Economics, 118(3), pp. 843-77.

[17] Kane, T.J., and D.O. Staiger. (2008). Estimating Teacher Impacts on Stu- dent Achieve- ment: An

Experimental Evaluation.. National Bureau of Eco- nomic Research Working Paper 14607.

[18] Murnane, R. J. (1975). The Impact of School Resources on Learning of Inner City Children. Cambridge, MA: Ballinger Publishing.

[19] Riehl, E., and Saavedra, Juan E. Urquiola, M. (2017). Learning and Earn- ing: An Approximation to College Value Added in Two Dimensions. NBER Chapters, in: Productivity in Higher Education National Bureau of Eco- nomic Research, Inc.

[20] Rivkin, S. G., Eric A. Hanushek, and John F. Kain. (2005). Teachers, Schools, and Academic Achievement.. Econometrica 73 (2): 41758.

[21] Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. American Economic Review 94 (2): 24752.