Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione

WORKING PAPER N. 43/2019

INVALSI data for school system improvement: the value added

Michele Cardone - INVALSI

Patrizia Falzetti - INVALSI

Chiara Sacco - INVALSI

Collana: Working Papers INVALSI

ISSN: 2611 - 5719

Abstract

INVALSI tests are carried since 10 years on solid scientific basis, they involved classes of five school-grades and produce system benchmarks and single student results deepen with several statistics. These data are firstly returned to all schools through a personal web area, then to other stakeholders such as the Minister of Education, regional representatives, researchers and the public community in general. INVALSI improves reliability and correct use of all these data using statistics procedures and providing help tools for the right use and interpretation. In particular, INVALSI returned to schools the school value added indicator, that is computed using a complex methodology but it is easily understandable for schools. Indeed, the school value added together with the observed score in INVALSI standardized test provide a useful tool for direct the school in a process improvement plan.

Parole chiave: big data, miglioramento scolastico, valutazione, modello multilevel, valore aggiunto

Keywords: big data, school improvement, evaluation, multilevel model, value added

Introduction

National Institute for the Evaluation of the System of Education and Training (INVALSI) conducts every year a large-scale survey to measure students' skills in Italian language and Mathematical knowledge and to provide an image of the school system's progress. INVALSI provides data to different users (Table 1): the official ministerial evaluation, stakeholders and public community, schools and researchers. Since the school autonomy reform, the necessity of an evaluation system based on standardized national test to evaluate student's performance has been proven to be a key element to identify the system's problems (Cipollone e Sestito, 2012; Siniscalco, 2012; Grimaldi and Serpieri, 2013).

At the beginning of every new school year, INVALSI returns to each school its results of all the classes involved in the survey, to provide tools for school improvement and for a self-evaluation in a comparative perspective. In spite of the intense debate surrounding several methodological concerns about the estimation of the school efficacy (Leckie and Goldstein, 2009; Koedel et al, 2014; Guarino et al, 2015), value added models still result to be useful to analyze the effectiveness of a school system (Thomas, 2001; Scheerens, 2011; Ferrao and Couto, 2014; Page et al. 2016)

The aim of this paper is to illustrate the procedure proposed by INVALSI for the estimation of school value added and to show how the elaborations conducted on this data generate an easy and usable information.

Table 1: Tools provided by INVALSI

Tool	Subjects	Aims	Data detail published	Deepest data detail provided	Periodicity
INVALSI school data area	- school subjects - other stakeholders - regional representatives	self-evaluation and improvement - at class level - at school level - at regional level	school, class	single students	annual (since 2012-13)
INVALSI public data area	- public (on line pre-registration)	to share all INVALSI databases	-	single students (anonymous)	annual (since 2004-05)
'Results report' and 'Technical report'	- public community - policy makers	institutional overview of the school system (based on a sample[1])	national, regional	-	annual (since 2007-08)
RAV Indicators	- MIUR (Ministry of Education) - autonomous provinces and regions - all single regions when requested	- MIUR data platform 'Scuola in Chiaro' - official self-evaluation school reports - regional assessments	school, class	class	annual (since 2013-14)

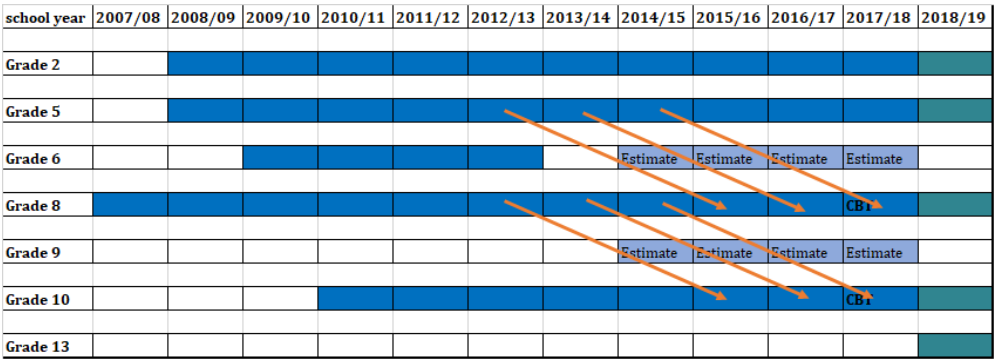
[1] The sample is conducted for each school grade involved and is representative at regional level.

Data provided for schools: school and class details

The first large scale survey was in 2007/08, consisting of a test at the end of the 8th grade which was included in the final school cycle exam (Quintano et al, 2010). Currently, as shown in blue in Figure 1, INVALSI every year involves 2nd, 5th, 8th and 10th grade students and, from this year, the 13rd grade in Italian language, Mathematics and English (Reading and Listening) standardized tests. All the published data (used for the annual reports[2]) and territorial benchmarks are computed on a school sample survey, which follow an administration protocol realized to ensure data reliability. In 2017-2018, INVALSI introduced the Computer Based Tests (CBT) for the 8th and 10th grades, and some new features are in progress like a new test for the 13th grade students, which is the end of the last school cycle (INVALSI, 2018).

INVALSI data involved on May 2018 more than 100.000 classes of four school grades and are returned to each school around September/October of the same year through their personal area available on our website[3]. This platform is accessible by personal login: in a unified web portal all different school profiles (from school manager to single teachers) can dispose of national and international assessment data[4] of the last four school years, with tutorials for the correct interpretation of data and figures.

Figure 1. National assessment history



Data returned to schools cover mainly these aspects:

- the general learning performances for all classes and the whole school, compared with territorial benchmarks (regional, macro regional, national) and with the average performances of 'similar' classes/schools comparable on the basis of social-economic and cultural context[5];

[2] http://www.invalsi.it/invalsi/doc_evidenza/2018/Rapporto_prove_INVALSI_2018.pdf
[3] http://invalsi-dati.cineca.it/
[4] Pisa, Timss.
[5] Only at 5th, 8th and 10th grade, where the student questionnaire is administered.

- variance decomposition of performances and social-economic and cultural context (ESCS) indicator between/within classes;
- student's distribution by learning levels, according to the test performance;
- microdata (structural data, Italian language and Maths answers and scores, for each single student);
- school value added.

School value added estimation

The school value added is an indicator of the school contribution to student learnings after controlling for the individual-specific variability among students (Raundebush and Willms, 1995). The school value added is estimated using a multilevel model (Bryk and Raudenbush, 2002; Goldstein, 2011) in order to account for the hierarchical structure of the data, in which students are nested within school. Thus, $\boldsymbol{Y_{ij}}$, the academic performance of the student *i* of the school *j*, is modelled as followed:

$$\boldsymbol{Y_{ij} = \beta_0 + \beta X_{ij} + u_j + \epsilon_{ij}}$$

where $\boldsymbol{\beta_0}$ is the intercept of the model, $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed-effects parameters, $\boldsymbol{X_{ij}}$ represents the vector of p predictors for the student *i* of the school *j*, $\boldsymbol{u_j}$ is the random effect associated with the school *j* with mean zero and variance $\boldsymbol{\sigma_u^2}$ and $\boldsymbol{\epsilon_{ij}}$ is the individual normal error with mean zero and variance $\boldsymbol{\sigma_\varepsilon^2}$. Thus, a multilevel model is a generalization of the linear regression model that allow to analyze data generated from several sources of variation instead of just one. Indeed, in the model shown in the previous equation, there are two sources of random variance: the random variance between the schools and the random variance among the students within the schools.

The term $\boldsymbol{u_j}$, estimated by the residual component of the multilevel model, allows to model the dependence among the students that are attending the same school and captures the school value added, an unobservable quantity that characterizes the school *j* and is shared by all its students.

For the estimation of the school value added, a fundamental element is the inclusion as predictors of only the exogenous variables, that are those variables which value is not determined in the system being studied. The Table 2 reports the list of the variables, both at student and school level, included in the model. Since $\boldsymbol{X_{ij}}$ predictors have both within-group and between-group variability (Snijders and Bosker, 1994), in order to

separate this two components from the total variation, an orthogonal reparametrization of the model has been performed (Paccagnella, 2006): level-student variables have been centered with respect to group mean.

Table 2. List of variables used to estimate the school value added.

A – STUDENT-LEVEL VARIABLES
1) Socio-demographic background
• Index of social, economical and cultural status of the student's family • Gender (indicator variable for female) • First-generation immigrants • Second-generation immigrants • Primary language spoken at home (indicator variable for whether Italian is not the language spoken at home) • Regularly dialect speakers (indicator variable for regular use of dialect at home)
2) Scholastic profile
• Prior score at INVALSI test • Pre-primary school attendance • Repeating student (indicator variable for student tha have repeated a grade at least once during their compulsory schooling
B – SCHOOL-LEVEL VARIABLES
Context
• Mean of students' score obtained at the previous INVALSI test • Index of social, economical and cultural status of the student's family • Percentage of first-generation and second-generation immigrants • Percentage of students that have repeated a grade at least once during their compulsory schooling • Number of enrolled students

INVALSI returns to schools the estimation of the school value added categorized in five categories using as threshold the national (or regional) mean of the school value added plus or minus 1 or 2 standard deviations. After controlling for the individual-specific variability among students, the school effect is categorized in the following classes:

1. *Negative school effect*: removing from the student performance the effect of the exogenous variables, the efficacy of the school is lower than the national/regional mean;
2. *Slightly negative school effect*: removing from the student performance the effect of the exogenous variables, the efficacy of the school is slightly lower than the national/regional mean;
3. *School effect equal to the mean*: the differences between the students' academic performance are explained by exogenous characteristics and the school efficacy is equal to the national/regional mean;

4. *Slightly positive school effect*: removing from the student performance the effect of the exogenous variables, the efficacy of the school is slightly higher than the national/regional mean;
5. *Positive school effect:* removing from the student performance the effect of the exogenous variables, the efficacy of the school is higher than the national/regional mean.

An example of distribution of school effect about grade 5 and grade 8 is reported at national level in Table 3 and Table 4; it is more interesting to observe the same effect in different Italian geographical areas (Table 5 and Table 6) where it is possible to note that the South has a very difficult situation especially in Grade 8 with the most percentage of schools in the negative class of value added.

Table 3. National value added distribution - Grade 5

School value added at national level - Grade 5					
	Positive	**Slightly positive**	**In the average**	**Slightly negative**	**Negative**
Italian	1,8% (81)	11,6% (516)	73,2% (3.258)	10,4% (463)	2,9% (131)
Mathematics	3,2% (144)	13,2% (588)	73,6% (3274)	12,0% (534)	2,3% (103)

Table 4. National value added distribution - Grade 8

School value added at national level - Grade 8					
	Positive	**Slightly positive**	**On the average**	**Slightly negative**	**Negative**
Italian	2,0% (76)	15,2% (581)	69,1% (2633)	10,8% (410)	3,0% (113)
Mathematics	2,2% (88)	15,2% (579)	72,1% (2751)	11,6% (444)	2,6% (100)

Tabel 5. National value added distribution between geographical areas - Grade 5

	Positive	**Slightly positive**	**On the average**	**Slightly negative**	**Negative**
North	1,2%	9,7%	79,7%	8,3%	1,2%
Centre	2,9%	15,1%	70,9%	9,4%	1,8%
South	4,5%	14,5%	59,0%	16,6%	5,5%

Tabel 6. National value added distribution between geographical areas - Grade 8

	Positive	**Slightly positive**	**On the average**	**Slightly negative**	**Negative**
North	2,2%	21,0%	71,9%	4,4%	0,4%
Centre	1,4%	9,9%	74,9%	11,6%	2,2%
South	2,3%	5,6%	60,0%	24,1%	8,0%

This procedure allows to obtain an indicator of the school efficacy that is modelled starting from the whole INVALSI data but can be easily interpreted and is usable by each single school. The indicator permits to give

at each school a measure of the school self-efficacy respect to geographical benchmarks. Moreover, the school value added together with the observed score in INVALSI standardized test (Figure 2) allows schools to identify their weaknesses for planning improvement actions (Ponisciack and Bryk, 2005; Rothman, 2010).

Conclusions

INVALSI data assessments gained a scientific and institutional recognition and became an increasingly common tool adopted by schools at local and institutional level. For example, they are used to guide parents' school choices and attract registrations or as a qualification to participate at local projects or at specific regional or national interventions.

In the last decades the accountability, responsiveness and self-improvement of schools have gained more interest in the research field (Sheerens, 2000a, Sheerens, 2000b, Ellet and Teddlie, 2003). The research activity in the evaluation of the educational institutions collects an increasing interest (Goldstein et al, 2000, Sestito, 2012, Harris, 2011); in particular, the necessity to be linked to real applications in school have promoted a debate to understand how to rank schools and how to measure school effectiveness (OECD, 2008; Martini and Ricci, 2010; Rosa, 2013; Rosa and Silva, 2014). The effectiveness is an indicator of the degree of achievement of its institutional targets. INVALSI provides tools to help schools to identify its strengths and weaknesses in order to encourage the planning of improvement actions and to promote a self-evaluation in a comparative perspective. The value added indicator, that every year INVALSI releases to each secondary school in Italy, measures the school effectiveness. The estimation of the value added is computed in order to provide the effectiveness of a single school in a comparative setting; in other words, the value added is a measure of the performance adjusted for all the factors that are out of the control of the institution, as for example the socio-economic and cultural status of the student's family, the gender of the students, the language spoken at home, etc. In this way, the effectiveness is completely due to the behavior of the institution itself. To help the school in the interpretation of the indicator, after the estimation using the linear mixed model, the value added is divided into classes that allow identifying the performance of school with respect to the nation or the reference region. Using this categorization, we achieved two different fundamental goals: the computation of an indicator that is easily understandable for schools and the provision of a tool that, with immediacy, allows comparing the school effectiveness with the territorial benchmarks.

Figure 2: Cross table observed score versus school value added

	Positive school effect	**Slightly positive school effect**	**School effect equal to the average**	**Slightly negative school effect**	**Negative school effect**
Observed score above the average	School contribution very evident Results good	School contribution evident Results good	School contribution in the average Results good	School contribution not adequate Results good	School contribution inadequate Results good
Observed score equal to the average	School contribution very evident Results acceptable	School contribution evident Results acceptable	School contribution in the average Results acceptable	School contribution not adequate Results acceptable	School contribution inadequate Results acceptable
Observed score below the average	School contribution very evident Results to be improved	School contribution evident Results to be improved	School contribution in the average Results to be improved	School contribution not adequate Results to be improved	School contribution inadequate Results to be improved

Reference

Cipollone, P., & Sestito, P. (2010). *Il capitale umano*. Il mulino.

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of personnel evaluation in education, 17*(1), 101-128.

Ferrao, M. E. and Couto, A. P. (2014) The use of a school value-added model for educational improvement: a case study from the Portuguese primary education system. *School Effectiveness and School Improvement*, 25, 174–190.

Goldstein, H. (2011). *Multilevel statistical models* (Vol. 922). John Wiley & Sons.

Goldstein, H., Huiqi, P., Rath, T., & Hill, N. (2000). *The use of value added information in judging school performance*. London: Institute of Education, University of London.

Grimaldi, E., & Serpieri, R. (2013). Privatising education policy-making in Italy: New governance and the reculturing of a welfarist education state. *Education Inquiry*, 4 (3), 22615.

Guarino, C., Reckase, M., Stacy, B. and Wooldridge, J. (2015) A comparison of student growth percentile and value-added models of teacher performance. *Statistics and Public Policy*, 2 (1), 1-11.

Harris, D. N. (2011). *Value-Added Measures in Education: What Every Educator Needs to Know*. Harvard Education Press. 8 Story Street First Floor, Cambridge, MA 02138.

INVALSI (2018), *Rilevazioni Nazionali degli apprendimenti 2017-2018. Rapporto Risultati*, INVALSI (http://www.invalsi.it/invalsi/doc_evidenza/2018/Rapporto_prove_INVALSI_2018.pdf)

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47, 180-195.

Leckie, G. and Goldstein, H. (2009) The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society A*, 172, 835–851.

Martini, A., & Ricci, R. (2010). Un esperimento di misurazione del valore aggiunto delle scuole sulla base dei dati PISA 2006 del Veneto. *Rivista di Economia e Statistica del territorio*.

OECD. Publishing, & Centre for Educational Research and Innovation. (2008). *Measuring improvements in learning outcomes: Best practices to assess the value-added of schools*. Organisation for Economic Co-operation and Development.

Paccagnella, O. (2006). Centering or not centering in multilevel models? The role of the group mean and the assessment of group effects. *Evaluation review*, 30(1), 66-85.

Page, G. L., Martín, E. S., Orellana, J., & González, J. (2017). Exploring complete school effectiveness via quantile value added. *Journal of the Royal Statistical Society A*,180(1), 315-340.

Ponisciak, S. M., & Bryk, A. S. (2005). Value-added analysis of the Chicago public schools: An application of hierarchical models. *Value-added modeling: Issues with theory and applications*, 40-81.

Quintano, C., Castellano, R., & Longobardi, S. (2010). La lunga e difficile prospettiva dell'adozione in Italia di valutazioni scolastiche standardizzate, Aspetti e problemi in riferimento alle esperienze statunitense e inglese. *Le nuove frontiere della SCUOLA*, 24, 62-71.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1), Sage.

Raudenbush, S. W., & Willms, J. (1995). The estimation of school effects. *Journal of educational and behavioral statistics*, 20(4), 307-335.

Rosa, A. (2013). *Il valore aggiunto come misura di efficacia scolastica. Un'indagine empirica nella scuola secondaria di primo grado* (Vol. 5). Edizioni Nuova Cultura.

Rosa, A., & Silva, L. (2014). Uno studio longitudinale sul valore aggiunto come misura di efficacia scolastica: risultati ed elementi di problematicità. *Italian Journal of Educational Research*, (12), 169-184.

Rothman, R. (2010). Beyond Test Scores: Adding Value to Assessment. *School Administrator*, 67(2), 20-24.

Scheerens, J. (2000a). *Improving school effectiveness* (Vol. 68). Paris: UNESCO, International Institute for educational planning.

Scheerens, J. (2000b). Autovalutazione e uso delle informazioni nella scuola orientata ai risultati. G. Barzanò, S. Mosca. e Scheerens, J.(a cura di), *L'autovalutazione nella scuola*. Milano: Bruno Mondatori.

Scheerens, J. (2011) Measuring teaching using value-added modeling: the imperfect panacea. *NASSP Bulletin*, 95, 122–140.

Sestito, P. (2012). La valutazione del valore aggiunto della scuola. *Italian Journal of Educational Research*, 27-36.

Siniscalco, M. T. (2013). Dieci anni di OCSE-PISA in Italia. N. Bottani, D. Checchi (a cura di), *La sfida della valutazione*, il Mulino, Bologna, 65-96.

Snijders, T. A., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological methods & research*, 22(3), 342-363.

Thomas, S. (2001). Dimensions of secondary school effectiveness: Comparative analyses across regions. *School Effectiveness and School Improvement*, *12*(3), 285-322.