Assessing the predictive capability of Invalsi tests on high school final mark

Silvia Bacci, Bruno Bertaccini, Alessandra Petrucci, Valentina Tocchioni

1. Introduction

Educational achievement can be considered a multifaceted issue, which takes into account many domains of learning at different levels of the educational path. In Italy, during the secondary school years, such achievements are measured through the administration of the INVALSI tests, which are standardized tests on a national scale that students carry out at different stages of their career, to identify their level of competence in subjects like literacy, numeracy, and English reading and listening proficiencies. They are applied each year to trace a history of students' skills and knowledge, but also to assess the correspondence between skills and competences acquired with respect to ministerial educational programs. Moreover, the high school final mark may be considered an overall result of performance at the end of secondary school, a sort of synthesis of several achievements and marks in different subjects.

The aim of the present work is to discover if and how the INVALSI scores and the high school final marks are related. More specifically, we intend to verify how the INVALSI scores are associated with students' high school final mark, taking into account students' characteristics as well as school observed (mainly, type of high school) and unobservable characteristics.

The present contribution represents a preparatory work to analyse the predictive capability of INVALSI scores and/or high school final marks on university students' careers. For this reason, the analysis is carried on the INVALSI dataset related with students enrolled in an Italian university.

In the next section, we describe data and statistical methods used in the study. Then, we illustrate the main results. A preliminary discussion of results and some final remarks about future research conclude the work.

2. Data and methods

To analyse university students' career in light of their performances during high school we use MobySU.it, a database that integrates multiple data sources, such as the Anagrafe Nazionale Studenti (ANS) data file, the INVALSI data file and the High School database. ANS is a government administrative database on the population of students enrolled in an Italian university between 2010 and 2020. The ANS data contain information on university students' career, individual characteristics, and high school background. The INVALSI data collect information on high school students' performances who obtained the high school diploma in 2019 and 2020. For each student, the following information are available: Economic and Social Status indicator (ESCS), students' INVALSI test scores in English (reading and listening), Italian, and Maths for grades 10^{th} and 13^{th} (i.e., high school second and fifth year), parents' education and type of employment, as well as other information about school, class and the student him/herself. These two sources of information at the student-level are merged using exact matching. Finally, the High School database includes aggregate data on all Italian high schools between 2015 and 2020, providing information on school characteristics (e.g., geographical area in which it is located, type of released degrees, and so on) and the number of students (grouped by gender) admitted to the final exam and of those who got the diploma.

We select 194,778 students who obtained the high school diploma in Italy in 2019 and enrolled

Silvia Bacci, University of Florence, Italy, silvia.bacci@unifi.it, 0000-0001-8097-3870
Bruno Bertaccini, University of Florence, Italy, bruno.bertaccini@unifi.it, 0000-0002-5816-2964
Alessandra Petrucci, University of Florence, Italy, alessandra.petrucci@unifi.it, 0000-0001-9952-0396
Valentina Tocchioni, University of Florence, Italy, valentina.tocchioni@unifi.it, 0000-0002-0793-6122

Referee List (DOI 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Silvia Bacci, Bruno Bertaccini, Alessandra Petrucci, Valentina Tocchioni, *Assessing the predictive capability of Invalsi tests on high school final mark*, ©Author(s), CC BY 4.0, DOI 10.36253/979-12-215-0106-3.03, in Enrico di Bella, Luigi Fabbris, Corrado Lagazio (edited by), ASA 2022 Data-Driven Decision Making. Book of short papers, pp. 11-16, 2023, published by Firenze University Press and Genova University Press, ISBN 979-12-215-0106-3, DOI 10.36253/979-12-215-0106-3

in an Italian university in the academic year 2019/2020. To verify if and how INVALSI scores are associated with students' high school final mark, we estimate a random intercept proportional odds model (Goldstein, 2010; Liu and Agresti, 2005; Snjiders and Bosker, 2012) with students as lower-level units and high schools as upper-level units, formulated as follows:

$$logit\left[P\left(Y_{ij} > y_c | X_{ij}\right)\right] = \beta X_{ij} + \gamma Z_j + u_j - \alpha_c$$

with i the generic student (i = 1, …, 194778), j the high school (j = 1, ..., 5203), and c = 1, 2, 3, 4 the four thresholds corresponding to the five categories in which the students' high school final mark was classified. As currently the high school final mark in Italy ranges from 60 to 100 cum laude, the response variable of the model was constructed by defining five ordinal categories: categories 1 to 4 represent 10 points of the high school mark range (i.e., 60-69, 70-79, 80-89, 90-99) and category 5 collects together 100 and 100 cum laude. Moreover, β and γ denote the vectors of regression coefficients for individual and school-level covariates, X_{ij} and Z_j respectively; u_j is the random intercept capturing the unobserved heterogeneity due to unobservable differences among schools, and α_c is a response category-specific threshold parameter. The random effects u_j are assumed to be normally distributed, with mean 0 and constant variance.

The explanatory variables of primary interest are the four students' INVALSI scores on English (reading and listening), Italian, and Maths on grade 13^{th}, and are included as standardized, continuous variables. The effect of INVALSI scores is controlled for both students' and schools' characteristics: at student-level we consider the student's gender, citizenship (Italian or not) and the student's macroarea of residence (North, Centre, and South/Islands); at school-level we consider the type of high school management (public vs. private), the type of high school attended (classified in seven categories: see Table 2 below), the percentage of high school graduates older than expected age at graduation, and the average ESCS of the school.

3. Results

Table 1 shows the median and mean values obtained in INVALSI scores by female students and male students, respectively, and predicted probabilities of obtaining one of the five mark categories for a median, female student and a median, male student[1]. The median female student has the highest probability (nearly 4 out of 10 students) of obtaining a high school final mark between 70 and 79 points, whereas the median male student has the highest probability (more than 1 out of 2 students) of obtaining a high school final mark between 60 and 69 points, namely the lowest category. Despite both groups have a low probability of obtaining a score equal to 90 or above, female median students seem to obtain higher scores than their male counterparts.

Table 2 shows predicted probabilities of a high school final mark between 60 and 69 points and between 100 and 100 cum laude, for a female/male student that obtained extreme scores to the INVALSI test, namely equal to the 10^{th} percentile and to the 90^{th} percentile in all four INVALSI scores (other control variables were set at the reference value), by type of high school. On one hand, predicted probabilities of a low final mark (60-69) are very high for those students who obtained an INVALSI score at the 10^{th} percentile. This result is confirmed throughout the different types of schools and for both genders, confirming how low scores on INVALSI tests are associated with low high school final marks. Nevertheless, students from vocational institutes, especially female students, report lower predicted probabilities, thus suggesting that these types of school may tend to give higher final scores than other schools, on average. Moreover, predicted probabilities of a low final mark are always higher for male students than for female students across all schools, thus suggesting that female students outperform male students.

[1] A median student is an Italian student that lives in the North of Italy, obtained a median score in the four INVALSI tests, and attended a scientific high school with a median percentage of high school graduates older than expected and a median ESCS at the school level.

Table 1: Median and mean values of INVALSI scores and predicted probabilities of obtaining a high school final mark for the median profile by gender.

	Female student	Male student
Median value (mean value)		
INVALSI score on Italian	216.8 (216.5)	218.2 (216.9)
INVALSI score on Maths	207.9 (209.0)	229.4 (228.4)
INVALSI score on English reading	220.4 (217.5)	222.6 (219.0)
INVALSI score on English listening	214.8 (214.0)	217.1 (216.2)
Predicted probability		
Pr(60-69 score)	0.359	0.509
Pr(70-79 score)	0.388	0.338
Pr(80-89 score)	0.159	0.101
Pr(90-99 score)	0.071	0.039
Pr(100-100L score)	0.024	0.012

Table 2: Predicted probabilities of high school final mark categories, by gender and type of high school. Extreme profiles (10th/90th percentile of INVALSI scores)

	Pr(60-69 score)	Pr(100-100L score)	Pr(60-69 score)	Pr(100-100L score)
	Scientific high school		***Classical high school***	
10percentile F	0.852	0.002	0.713	0.005
10percentile M	0.917	0.001	0.823	0.002
90percentile F	0.054	0.203	0.023	0.368
90percentile M	0.099	0.118	0.044	0.238
	Applied sciences high school		***Foreign language high school***	
10percentile F	0.861	0.002	0.723	0.004
10percentile M	0.922	0.001	0.831	0.002
90percentile F	0.058	0.191	0.024	0.357
90percentile M	0.106	0.110	0.046	0.229
	Technical institute		***Vocational institute***	
10percentile F	0.628	0.007	0.399	0.020
10percentile M	0.759	0.004	0.551	0.010
90percentile F	0.015	0.460	0.005	0.685
90percentile M	0.029	0.315	0.011	0.540
	Other high school			
10percentile F	0.619	0.007		
10percentile M	0.751	0.004		
90percentile F	0.014	0.470		
90percentile M	0.028	0.324		
Sample size	***194,778***			

Note: other covariates are set at the reference value/mean value.

On the other hand, INVALSI scores at the 90th percentile tend to be associated with high final marks (100 and 100 cum laude), with differences varying according to the type of school. More precisely, students who attended scientific high schools and applied sciences high schools report predicted probabilities lower than 0.25, whereas students who attended technical institutes and vocational institutes show probabilities of high final marks definitely higher. This result outlines the presence of a significant interaction effect between type of high school and INVALSI score on the high school final mark. Coherently with low final marks, predicted probabilities of a high final mark are always higher for female students than for male students across all schools, thus suggesting again that female students outperform male students.

Finally, coherently with a positive association between INVALSI scores and high school final mark, predicted probabilities of having a high final mark are very unlikely for those students who obtained an INVALSI score at the 10th percentile, as well as predicted probabilities of having a low final mark are unlikely for those students who obtained an INVALSI score at the 90th percentile.

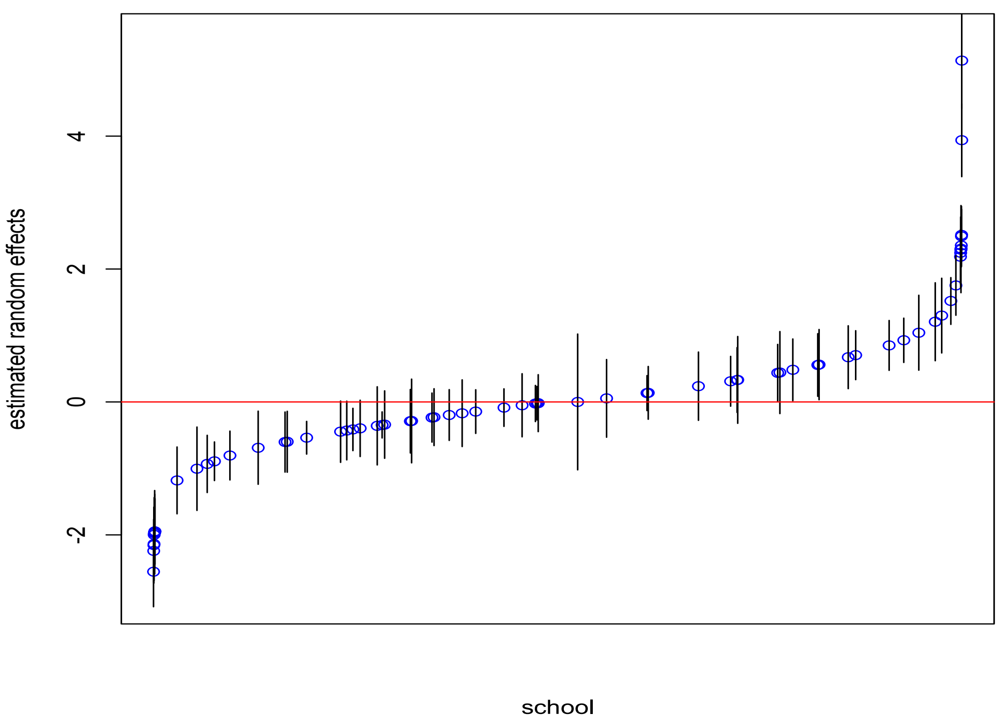
Lastly, Table 3 shows the estimated coefficients for all covariates included in the models. To sum up the effect of variables on high school final mark, all INVALSI scores are positively associated with the school final mark, as well as female students (with respect to male students), residing in the Centre and South of Italy (instead of residing in the North), attending a private school (in comparison with a public school) have all a positive effect on the likelihood of a high final mark. Conversely, being a foreign student has a negative effect on a high final mark. As for the type of high school, all schools have a positive effect on the likelihood of a high final mark with respect to students attending a scientific high school, except students attending an applied science high school, whose coefficient is negative (but only slightly significant). Finally, the two second-level covariates appear to be significant: indeed, both the high school ESCS and the percentage of graduates over 19 in the high school have a negative association with the high school final mark.

Table 3: Model coefficients for the multilevel proportional odds model on high school final mark categories (Sample size: 194,778).

	Coeff.	**SE**	**P-value**
INVALSI score on Italian	0.559	0.007	0.000
INVALSI score on Maths	0.856	0.007	0.000
INVALSI score on English reading	0.243	0.007	0.000
INVALSI score on English listening	0.296	0.007	0.000
Gender (ref. Male)			
Female	0.689	0.010	0.000
Citizenship (ref. Italian)			
Foreign	-0.343	0.024	0.000
Macroarea of residence (ref. North)			
Centre	1.039	0.032	0.000
South	1.974	0.029	0.000
School property (ref. Public)			
Private	0.455	0.053	0.000
Type of high school (ref. Scientific high school)			
Classical high school	0.913	0.044	0.000
Applied sciences high school	-0.081	0.042	0.053
Foreign language high school	0.857	0.042	0.000
Other high school	1.384	0.042	0.000
Technical institute	1.339	0.043	0.000
Vocational institute	2.385	0.073	0.000
High school ESCS	-0.424	0.039	0.000
% graduates over 19 in high school	-0.004	0.001	0.001
Thresholds			
First: 60-69 score	-0.065	0.037	
Second: 70-79 score	1.786	0.037	
Third: 80-89 score	3.049	0.037	
Fourth: 90-99 score	4.556	0.038	
Random part			
Variance at the high school level	0.542	0.014	

Finally, from Table 3 we observe that the school-level variance is statistically significant and represents the 14% (intraclass correlation coefficient) of the total variance of the response variable explained by the hierarchical structure of data. In more detail, the estimated school-level random effects are displayed in Figure 1 together with the related 95% confidence intervals. For ease of readability, the caterpillar plot reports only a sub-sample of schools: the ten schools with the lowest random effects (on the left side of the plot), the ten schools with the highest random effects (on the right side of the plot), and other fifty randomly selected schools (in the centre of the plot). It is worth to outline how schools at the extremes of the plot significantly differ from the other schools. Moreover, in the two extremes we found different schools (i.e. technical institutes such as classical and scientific high schools), as well as divers geographical location (i.e., Sicily, Tuscany, or Emilia-Romagna) without showing a precise pattern (for example, high schools with a positive influence are located both in South and in the Centre of Italy). At first sight, we could not find any systematic difference between high schools that may have a positive or negative influence on INVALSI scores, but a deeper interpretation is needed to check if potential differences exist.

Figure 1: Caterpillar plot: school-level estimated random effects with 95% confidence intervals for a sub-sample of schools.



4. Preliminary conclusions and future research

Our preliminary analyses show that the INVALSI scores are positively associated with the high school final mark, which may be considered an overall performance outcome at the end of the high school career, with higher INVALSI scores corresponding also to higher high school final marks. Despite it, some highlights are worth to be stressed. First, female students achieve high school final marks higher than male students, keeping constant the INVALSI scores and other characteristics. Second, differences by type of high schools are visible too, being constant the INVALSI scores and other characteristics. Third, the association between INVALSI scores and high school final marks seems to be stronger for lower scores/marks. These issues rise some doubts. On one side, they question about the real capability of INVALSI tests to predict the performance at the high school final examination; on the other side, the high school final evaluation is not exempt from disparities according to gender and type of school, irrespective the INVALSI scores.

Given these preliminary results, we will proceed with a deeper analysis of our results in the light of eventual differences on individual characteristics – such as student's geographical area of residence – and on school-level characteristics – such as high school quality (for example, in terms of percentage of graduates over 19). Moreover, in light of the discrepancies between INVALSI scores and high school final marks above outlined, both these types of information will be object of interest in a next step concerning the academic career of students in terms of credits earned at the first year of university. In particular, it will be of primary interest to investigate the predictive capability of INVALSI scores and the high school final mark, and the differences between them, also taking into account the high school of origin and the gender. More precisely, to analyse the predictive capability of the INVALSI scores and the high school final mark on the academic students' career (evaluated in terms of credits earned in the first year), we will estimate a multilevel model, to take into account that students are nested within athenaeums. Then, the functional form of the model will be chosen in accordance with the distribution of the number of credits earned in the first year, which, at first sight, does not seem distributed as a normal variable and shows one or two peaks around zero and/or sixty credits in most athenaeums. We will interpret our results in the light of assessing potential divergences in students' performances during the transition from high school to university.

References

Goldstein, H. (2010). *Multilevel Statistical Models*. 4th Edition, John Wiley & Sons, Ltd

Liu, I., Agresti, A. (2005). The analysis of ordered categorical data: An overview and a survey of recent developments. *Test*, **14**(1), pp. 1-73.

Snijders, T. A.B., Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London, Sage Publishers.