

## Risky decline? Exploring the determinants of pupil's proficiency development over time

Melisa L. Diaz Lema <sup>a,\*</sup> , Chiara Masci <sup>b</sup> , Mara Soncin <sup>a</sup> , Tommaso Agasisti <sup>a</sup> 

<sup>a</sup> Department of Management, Economics, and Industrial Engineering, Politecnico di Milano, Italy

<sup>b</sup> Department of Economics, Management, and Quantitative Methods, University of Milan, Italy

### ARTICLE INFO

#### Keywords:

At-risk students  
Proficiency development  
K-12 education  
MCMC estimation

### ABSTRACT

The early identification of at-risk students and of factors influencing proficiency development are crucial in education policy. This study investigates how students' fluctuations in academic performance during primary school impact the probability of being at-risk in Mathematics and Reading by the end of lower secondary school. Leveraging longitudinal data from 278,765 students in Italy, this study employs a multilevel multinomial model with Markov chain Monte Carlo (MCMC) techniques to explore personal, class, and school-related factors steering shifts between proficiency and non-proficiency. Findings highlight the importance of these transitions in predicting at-risk students at grade 8, emphasizing the role of peers' academic achievements and the variation across schools in shaping proficiency development.

### 1. Introduction

Longitudinal studies examining student progress have reported that those who are initially struggling academically in maths or reading are less likely to become highly proficient in subsequent years than their counterparts, and thus report a higher probability of becoming at-risk students [1]. Early detection of at-risk students—those who may face academic challenges and risk educational poverty—is a matter of primary interest to education policy [2]. Educational poverty refers to a lack of knowledge and/or skills that precludes participation in the labour market and possibly affects other spheres of life, such as family life and health, in ways not experienced by individuals with higher levels of education [3]. Identifying recurring personal, class, or school characteristics that can help them to increase their level of proficiency over time is crucial. Indeed, early education interventions prove to be more cost effective than remedial interventions proposed later in life [4].

The prompt identification of at-risk students has been proven effective in reducing student attrition rates [5]. Despite this, the bulk of research on at-risk students had centred around status measures—those describing the academic performance of a student at a single point in time—and high school students [6,7], neglecting the performance transitions or growth between primary and middle/lower secondary school, in which student disengagement often occurs [8]. This

constitutes a significant oversight, as status measures alone fail to account for student progress over time. A low-achieving student in an effective school may actually be improving, while a high-achieving student in an ineffective school may be declining. By focusing solely on static performance snapshots, status measures risk misidentifying schools' effectiveness and students' true academic trajectories [9]. Additionally, the earlier the interventions are provided to at-risk students, the more effective they are likely to be [10], especially since improved performance in specific abilities over time has a favourable impact on their future academic success [11].

Proficiency development represents the progression required to either meet or fall short of a standard level of proficiency [12]. Hence, a certain degree of academic growth may culminate in either proficiency or non-proficiency. In this study, non-proficient students are categorized as at-risk students (and we will use these terms interchangeably).

Policymakers have increasingly acknowledged the importance of assessing student growth rather than solely relying on single achievement scores, particularly in the design of accountability systems [12, 13]. In Italy, descriptive proficiency levels for the last year of lower secondary school (grade 8) in mathematics and reading are provided by INVALSI, the National Evaluation Committee that yearly test Italian students at given grades. While students are tested also in grade 2 and 5 (see Section III for further details), grade 8 is the first in which student's

\* Corresponding author.

E-mail addresses: [melisalucia.diaz@polimi.it](mailto:melisalucia.diaz@polimi.it) (M.L. Diaz Lema), [chiara.masci@unimi.it](mailto:chiara.masci@unimi.it) (C. Masci), [mara.soncin@polimi.it](mailto:mara.soncin@polimi.it) (M. Soncin), [tommaso.agasisti@polimi.it](mailto:tommaso.agasisti@polimi.it) (T. Agasisti).

<https://doi.org/10.1016/j.seps.2025.102207>

Received 9 February 2024; Received in revised form 18 February 2025; Accepted 18 March 2025

Available online 20 March 2025

0038-0121/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

proficiency in the test is measured by these levels (in addition to test score), which are developed with reference to the method used in the PISA survey [14,15]. The levels are described in terms of what students at a certain level typically, and with a certain degree of probability, know and can do in relation to the domain assessed in the test. This approach allows for better identification of students who do not reach a minimum level of proficiency in grade 8 (defined as reaching a level of 1 or 2 out of 5 possible levels). Analysing the longitudinal trajectories of these students can reveal the risk and protective factors influencing their future school performance.

In the broader literature, risk and protective factors in relation to future school performance can be found in various aspects of a child's life, including individual character, family situation and school and social environment [16,17]. While some studies had focused on the impact of students' personal traits on their academic growth [18], less is understood about the influence of school and class-related factors on overall student performance in general and proficiency development in particular [19]. To address these gaps in the literature, this study examines two research questions: (i) to what extent can proficiency development during primary school help to identify at-risk students in the final year of lower secondary school? (ii) What are the factors associated to positive and negative transitions in student performance during primary school? Focusing on the early stages of education is particularly important for policy purposes, as it can enable timely actions to be implemented to support at-risk students.

The paper makes two primary contributions to the field: first, it adds to the conceptual debate surrounding the prompt identification of at-risk students, who may face future academic difficulties; second, the study models proficiency development and identifies personal, class, and school-related factors that contribute to significant improvements or declines in student performance during primary school, adding a time-variant dimension to a discussion which is usually conducted in a static manner.

The paper is structured as follows: section 2 provides a comprehensive literature review on at-risk students and factors affecting academic growth; section 3 describes the data and outlines the methods used in this study; section 4 presents the primary findings; and in Section 5, we present our conclusions and discuss policy implications.

## 2. Literature review

The academic literature defines at-risk students as those that are "are poorly equipped to perform up to academic standards" [20]. Being a student at-risk may result in a number of negative outcomes, including not only the absence of required skills, but also the failure to obtain formal qualifications. These consequences manifest at different time spans: in the short term, examples include grade repetition [21]; in the medium term school drop-out [22]; and in the long term, individuals may experience lower levels of income, health and well-being [23].

There is a considerable body of literature addressing the underlying causes or determinants of low educational performance (J.-W. [24]), sometimes highlighting relevant predictors for the prompt identification of non-proficient or at-risk students [7]. Emphasis is often placed on students' individual and family characteristics [25], especially in relation to socio-economically disadvantaged background [26]. Additional indicators such as absenteeism and unsatisfactory behaviours at school have been found to be highly predictive of at-risk students [27]. Nevertheless, the inclusion of these detailed variables in empirical studies is often restricted due to poor data availability [21] or cross-sectional predictors, where data for the predictor comes from a single academic year. That which is also often neglected is the role of school as a protective factor to prevent at-risk outcome [28]. Few studies reveal the relevant role of a positive relationship with teachers among these protective factors [29,30]. On the contrary, peer composition (intended as a larger proportion of peers coming from low socio-economic families) increases the likelihood of being

non-proficient on standardized tests [31].

Research has shown that a high percentage of at-risk students send distress signals early in their educational career [32], hence early identification and intervention can be highly beneficial for reducing later negative outcomes. Indeed, student failure during upper secondary school is often the result of a long-term process in which, rather than a single event happening in time, multiple stressors accumulate over time and eventually reach a threshold in which the student no longer attends school [7]. Capturing this time-changing process is important for accurately predicting early which students may be at-risk of not achieving the minimal competences and may dropout. Hence, utilizing longitudinal data that track students' changes over time, and exploring what produces such changes, can facilitate the identification of both risk and protective factors that can be targeted through interventions to improve outcomes for at-risk students. This research approach is still not very frequent though.

From a methodological perspective, multiple approaches exist for defining growth models. One such approach is the use of gain scores, which measure the differences between two or more test scores at different points in time aligned to a common scale. However, gain scores can be problematic, potentially resulting in distortions such as ceiling effects that restrict the gains of initially high-scoring students and spurious relationships between gains and initial status [12]. Student growth percentiles offer another approach by normatively situating achievement change, bypassing questions associated with the magnitude of change in favor of the relative amount of change [9]. However, this approach shows the same potential threats of gain scores.

A final alternative approach to creating standards for growth involves quantifying the level of growth necessary to reach prescribed levels of achievement and using those results to inform a standard setting procedure [9]. In this sense, target-based standard setting classifies students or groups as making adequate growth by determining if they are "on track" to meet a future target standard, such as reaching the proficiency cut point in a particular grade level [33]. This method ensures that growth adequacy criteria are student-specific, recognizing that students well below proficiency must demonstrate higher growth percentiles to reach proficiency compared to students near the threshold. Such nuanced approaches to growth models can provide a clearer picture of student development informing more effective interventions, and it is indeed the approach we are adopting along the study.

Among the studies adopting an early and longitudinal approach, Hamre & Pianta [34] followed 179 students from kindergarten to grade eight, showing the relevance of the impact of having a positive relationship with teachers to prevent negative academic and behavioural outcomes of students. However, their approach is more related to a long-term perspective than to time-varying observation of students' characteristics. Indeed, most of the studies addressing differences in academic achievement focused on the predictors of children's inter-individual differences [21,35], but little focus was placed on the determinants of children's intra-individual change specifically, on what predicts children's individual growth rates or proficiency development [11]. Goldstein [36] stressed the relevance of measuring change properly in longitudinal analyses. On a modelling perspective, growth models are typically adopted to account for multiple observations of students over time [37]. Shanley [38] built growth models for mathematics achievement between kindergarten and grade 8, finding a non-linear relation over time and a larger growth in the first years of schooling compared to the later stages.

The current paper contributes to the modelling of proficiency development on achievement, with a two-fold focus: (i) identifying the determinants of students' proficiency developments during primary school and (ii) investigating how these developments do influence the probability of being a student at-risk at secondary school. At both stages, we account for the across schools' variability, by measuring the school effect on the students' performance. To this end, varying from the

growth models statistical literature, we rely on multilevel models [39] for multinomial and binomial responses. While multilevel models for binary responses are broadly studied and applied in the statistical literature, the ones for unordered multinomial responses are still less investigated [40,41].

### 3. Data and methods

This paper's empirical analysis is based on data provided by the Italian National Evaluation Committee for Education (INVALSI), which administers standardized tests to all Italian students (census tests) in reading and mathematics to Italian students at specific grade levels each year, namely grade 2, 5 (primary school), 8 (lower secondary school), 10 and 13 (upper secondary school). Our focus is the test results for grade 8, the last year of lower secondary school, in the 2018/19 SY. These students took the INVALSI test in grade 5 (2015/16 SY) and 2 (2012/13 SY), making it possible to follow their proficiency development over time. In total, our sample comprises 278,765 students across 5446 schools. This sample comprises about 70 % of the students' cohort and the loss is mainly attributable to technical issues in tracking the student identifier over time. Specifically, cohort loss from grade 2 to grade 5 is nearly double (68,000 students missing) compared to the loss from grade 5 to grade 8 (32,000 missing values), indicating improvements in data collection and storage capabilities over time. A reduction due to student retention would be more threatening but it is less likely given this phenomenon is very reduced in lower secondary school – less than 2 % of students are retained [42] and it does not apply to primary school students.<sup>1</sup> Furthermore, a representativeness analysis confirmed that the characteristics of our sample are comparable to those of the entire population, as shown in [Appendix B](#).

Analyzing student proficiency at grade 8 is particularly interesting given that, from the test took at this grade onwards, INVALSI not only provides the student's score in the standardized test, but also defines the level of competence, in a range from 1 to 5, that the student reached in each subject of the test [43]. A competence level equal to 1 or 2 indicates that the student can identify basic information in text and solve easy mathematical problems but does not reach the minimum level of competence for that grade (level 3). These proficiency levels are determined through a psychometric analysis that assesses students' abilities, weighting each item according to the skills required to answer it. Level 3 corresponds to the minimum threshold of knowledge and skills that a student should have in a particular grade [44]. Building on this framework, this study adopts INVALSI's definition of an at-risk student, referring to those students who fail to meet INVALSI's minimum competence level for each subject [45]. Specifically, students with a competence level of 2 or lower (out of 5 possible levels) are considered at-risk. In our cohort, 34 % of students are at-risk in mathematics, while 29 % are at-risk in reading [Table 1].<sup>2</sup> We categorized covariates from the administrative dataset into four distinct groups for analysis. These are: (i) personal characteristics [25,26] and (ii) instruction and previous performance indicators [46], which are common factors aligning with findings from other empirical studies on at-risk students. There is also a growing perspective that frames the students' performances as a systemic failure (V. E. [47]). This conceptualization encompasses risk factors related to school characteristics and student transfers. Consequently, the remaining two categories of data in our study are derived from: (iii) primary school class characteristics and peers'

<sup>1</sup> Student retention is more prevalent in upper secondary school, where approximately 7 % of students are required to repeat the academic year (<https://www.istat.it/produzione-editoriale/annuario-statistico-italiano-2024>).

<sup>2</sup> To ensure a predictive setting, the available data was divided into a training set (70 %) and a test set (30 %). The models were developed using the training data, and their performance in identifying students at risk was evaluated exclusively on the test set.

performances and (iv) further environmental characteristics. [Table 1](#) reports the variables used in our analysis. The sample students are balanced in terms of gender, 92 % are native Italians and the majority (91 %) attended the kindergarten. Most of the students are regular, we observe a very low percentage of late and early enrolled students (1 % and 1 %, respectively) and 14 % of students change primary school between grade 2 and 5. In terms of class composition, the average number of students per class is 20 at grade 5 and the average socio-economic status (SES) is 0.07.

The empirical analysis is informed by an approach based on two steps as described below.

#### 3.1. Step 1: Predicting at-risk students at lower secondary school

The first phase of this study aims to explore the effect that proficiency transitions, or academic performance transitions during primary school, as measured by standardized tests from grade 2 to grade 5, have on students' future performance (at grade 8). Having access to information on INVALSI test scores and teacher-assigned marks for the same cohort of students when they were in grade 5 (SY 2015/2016) and grade 2 (SY 2012/2013), allowed us to identify and classify students based on their primary school transitions. The anonymous dataset was provided by INVALSI statistical office. We defined three categories of students based on their performance trajectory between grades 2 and 5: upgrading students (moving from the bottom 40 % at grade 2 to the top 40 % at grade 5 of the performance distribution), downgrading students (moving from the top 40 % at grade 2 to the bottom 40 % at grade 5), and unvarying students (showing no significant changes in performance during this period). The unvarying category collects all students who were neither upgrading nor downgrading, and thus showed non-significant performance variation over time. This group includes students whose performance remained in the same range of percentile of the distribution or moved from top/bottom to the middle or viceversa. The 40 % threshold has been chosen as the level of improvement or decline in primary school students' performance that best predicts students' level of risk in the final year of lower secondary school (grade 8). [Appendix A](#) reports the details of this analysis. About 15 % of students underwent a significant transition on academic performance during primary school, with 20,841 (7.5 %) upgrading students and 21,823 (7.8 %) downgrading students in reading, as well as 18,391 (6.6 %) upgrading and 19,262 (6.9 %) downgrading in mathematics [Table 2]. The identification of these categories is crucial for the purpose of the study, as it allows to observe developments between bottom/top performance distribution and to properly compare those students whose performance changed consistently over time, with those who did not (that can either be always top, middle or bottom performers). This categorization has the main implication to spotlight the student transition, in line with the rationale of the study.

The availability of a diverse range of individual and environmental characteristics spanning across multiple years offers us two distinct advantages. On one hand, we can compare the performance of logit models identifying at-risk students over time,<sup>3</sup> using covariates up to grade 5 and then using covariates up to grade 8. On the other hand, we can isolate the effects of other determinants of academic performance and focus specifically on how the interaction between the performance on grade 5 and the changes during primary school (whether upgrading, downgrading, or remaining consistent) impacts future outcomes in mathematics and reading at the end of lower secondary school. By considering these interactions, we account for both the performance at grade 5 and the changes leading up to it.

In particular, for each student  $i$ ,  $i = 1, \dots, N$ , our response variable  $y_i$  takes value 1 when the  $i$ -th student is at risk at grade 8 and 0 otherwise.

<sup>3</sup> Students between percentile 40 and 60 at grade 2 are excluded from the samples considered in the MCMC multi-response generalized model.

**Table 1**  
List and descriptive statistics of variables used in the empirical analysis.

Category	Variable	Type	N	Mean	Std. Dev.	Min	Max
Outcome	at-risk math	Binary	278765	0.34	0.47	0.00	1.00
	at-risk reading	Binary	278765	0.29	0.45	0.00	1.00
Personal characteristics	gender	Binary	278765	0.50	0.50	0.00	1.00
	Nationality	Cat	278765	100 %			
	-Italian		258111	92 %			
	-1st gen. immigrant		4427	2 %			
	-2nd gen. immigrant		16227	6 %			
Instruction and previous academic performance	Socio-economic status (ESCS index)	Num	278765	0.12	0.92	-2.85	2.34
	kindergarten attendance	Binary	278765	0.91	0.29	0.00	1.00
	enrollment	Cat	278765	100 %			
	-regular		272646	98 %			
	-early		3643	1 %			
	-late		2476	1 %			
	ta_marks_math_g8	Num	278765	6.93	1.39	1.00	10.00
	ta_marks_red_g8	Num	278765	7.11	1.16	1.00	10.00
	ta_marks_math_g5	Num	278765	8.01	1.06	1.00	10.00
	ta_marks_red_g5	Num	278765	7.94	1.01	1.00	10.00
	ta_marks_math_g2	Num	278765	8.19	1.00	1.00	10.00
	ta_marks_red_g2	Num	278765	8.08	1.02	1.00	10.00
	test_score_math_g5	Num	278765	0.41	1.20	-5.50	4.92
	test_score_red_g5	Num	278765	0.27	1.01	-5.73	4.03
	test_score_math_g2	Num	278765	0.51	1.26	-4.73	4.71
	test_score_red_g2	Num	278765	0.33	1.01	-4.47	3.80
Primary school class characteristics and peers' performances	ESCS_classe_g5	Num	278765	0.07	0.48	-2.30	1.83
	peer_imm1_g5	%	278765	0.01	0.03	0.00	0.67
	peer_imm2_g5	%	278765	0.05	0.07	0.00	0.88
	peer_fem_g5	%	278765	0.41	0.12	0.00	0.96
	n_stud_prev_g5	Num	278765	20.50	4.34	2.00	33.00
	delta_n_stud_prev	Num	278765	0.01	0.20	-0.91	12.00
	peer_score_g2_mat	Num	278765	0.41	0.85	-4.73	4.71
	peer_score_g5_mat	Num	278765	0.34	0.82	-2.83	4.92
	peer_score_g2_ita	Num	278765	0.25	0.63	-4.47	3.80
	peer_score_g5_ita	Num	278765	0.19	0.58	-2.68	4.03
Environment	reg. location	Cat	278765	100 %			
	-center		44145	16 %			
	-islands		28607	10 %			
	-northeast		55666	20 %			
	-northwest		74594	27 %			
	-south		75753	27 %			
	change_reg_g2g5	Binary	278765	0.006	0.08	0.00	1.00
	change_school_g2g5	Binary	278765	0.14	0.35	0.00	1.00
	change_reg_g5g8	Binary	278765	0.008	0.09	0.00	1.00
	change_school_g5g8	Binary	278765	0.39	0.49	0.00	1.00

**Table 2**  
Absolute and relative number of students for each transition category and each subject.

Subject	Always bottom	Upgrading	Middle	Downgrading	Always top
Mathematics	85,524 30.7 %	18,391 6.6 %	59,930 21.5 %	19,262 6.9 %	95,658 34.3 %
Reading	93,674 33.6 %	20,841 7.5 %	47,594 17.1 %	21,823 7.8 %	94,833 34.0 %

For each subject (mathematics and reading), the logistic regression takes the following form:

$$\text{logit}(y_i) = \sum_{p=0}^P \beta_p x_{ip} \tag{1}$$

where  $P$  is the total number of predictors (the  $P+1$  terms include the intercept). For each subject, we run two models that include the following covariates: student INVALSI test score at grade 5 and longitudinal changes (unvarying, downgrading, upgrading); students' personal characteristics (gender, SES, nationality); previous education and academic performance (kindergarten attendance, early/late enrolment, teacher-assigned marks - TA marks - at grade 5, change in TA marks between grade2 and 5); and environmental factors (class average SES, geographical area and a dummy for changing school or region). The first

model evaluates all these covariates at grade 5, while the second model evaluates them at grade 8 to test how much the predictive model gains when we consider variables that refer to the same time as the output (grade 8) [as shown in Table 3]. Regarding INVALSI test score at grade 5 and longitudinal changes, we consider their interaction as a combined predictor that estimates the impact of the performance at grade 5 depending on the type of longitudinal change. The choice of including this interaction among the covariates leads to a not trivial interpretation of results. Still, this choice stems from a specific interest in examining whether the impact of a given INVALSI score varies based on the student's performance history. For instance, we aim to explore whether a high/low INVALSI score has the same effect on a student who has consistently performed well/poorly in previous years compared to one who has shown significant improvement/worsening.

In this study, we adhere to the classical methodology of incorporating previous achievements (grades 2–5) to forecast future performance (grade 8) [48]. We recognize that this involves including an autoregressive term, thereby creating a dynamic model. Various methods for estimating such dynamic models have been proposed in the literature, offering strategies for obtaining consistent parameter estimates [49]. Given the categorical nature of our proficiency development measure (grades 2–5), an alternative approach could involve using the predicted probabilities of students' belonging to different profiles (e.g., upgrading, downgrading) rather than the profiles themselves. Being our primary interest to provide accurate predictions of grade 8 performance

**Table 3**

Covariates used in Step 1 logit models to predict at-risk students at grade 8, based on information available three years earlier (grade 5) and updated information at grade 8.

	Model - covariates grade 5	Model - covariates grade 8
ScoreG5: Longitudinal change (g5-g2)	ScoreG5: Unvarying ScoreG5: Downgrading ScoreG5: Upgrading	
Personal characteristics	Gender Student SES - ESCS Nationality	
Instruction and previous academic performance	TA marks – g5 Delta TA marks (g2 – g5) Kindergarten attendance Enrollment	TA marks – g8 Delta TA marks (g5 – g8)
Environment	Class SES – g5 Geo. macro-area Changing school G2-G5 same region Changing region G2-G5	Class SES – g8 Geo. macro-area Changing school G5-G8 same region Changing region G5-G8

Notes: Covariates listed in the middle of the table are included in both models (e.g., gender, student SES, and nationality). Covariates appearing only on the left are specific to grade 5, while those on the right are updated for grade 8 (e.g., teacher-assigned (TA) marks at grade 5 on the left and TA marks at grade 8 on the right).

and to examine the impact of primary school proficiency development on later outcomes, this represents an alternative approach that could be further explored.

### 3.2. Step 2: Characterizing students with proficiency transitions at primary school

The second part of the study identifies the students' characteristics as well as the class and school factors that affect students' proficiency development during primary school, pointing out the factors that differentiated the students who made progress towards proficiency from those who did not. In this second step, the definition of upgraded and downgraded student remains the same, while the unvarying classification needs further development. Given the vast heterogeneity of performance among the unvarying students [Fig. 1], we further disaggregate this category into multiple groups to make it a proper reference group. Thus, we differentiate within the unvarying group students who were always at the top of the distribution, students who were always at the bottom and students who always adhered to the middle of the distribution. As Table 2 shows, in Mathematics, 30.7 % of students consistently remained in the lower performance bracket, while 6.6 % (18,391 students) exhibited significant improvements in their performance. Conversely, 34.3 % consistently maintained their top-tier performance, with 6.9 % of students experiencing a decline.

### 3.3. MCMC multilevel multinomial model

To compare the effects of different covariates on upgrading the consistently bottom-performing students and downgrading the consistently top-performing students, this study uses a multilevel multinomial regression model that implements Markov Chain Monte Carlo (MCMC)

routines for fitting multi-response generalized linear mixed models [40].

In the perspective of characterizing the *varying* students compared with the *unvarying* ones, we consider two sub-samples of our dataset: (i) students who were at the bottom of the distribution at grade 2 (bottom 40 %) and (ii) students who were at the top of the distribution at grade 2 (top 40 %).<sup>4</sup> Students at the bottom (top) of the distribution at grade 2 might remain at the bottom (top) of the distribution at grade 5, upgrade (downgrade) to a medium level or upgrade to the top level (downgrade to the bottom level). For each student in the first (second) sample, we model the probability for him/her to upgrade to the top (downgrade to the bottom), upgrade (downgrade) to a medium level or to remain in the same portion of the distribution. The main idea here is to identify variables (at individual, class or school level) that are associated with the probability for a student to belong to the group of downgrading/upgrading one, by comparing him/her with a proper reference group (i.e. those remaining in their same initial performance group). Also, in order to take account of the heterogeneity at school level, we adopt a multilevel modelling, by considering students nested within primary schools, and we evaluate the effect that each school has on the longitudinal transitions of its students.

We build two models that regard the following response variables:

$$t_{ij} = \begin{cases} top & \pi_{top,ij} \\ middle & \pi_{middle,ij} \\ bottom & \pi_{bottom,ij} \end{cases} \quad b_{ij} = \begin{cases} top & \omega_{top,ij} \\ middle & \omega_{middle,ij} \\ bottom & \omega_{bottom,ij} \end{cases} \quad (2)$$

where  $t_{ij}$  (and, respectively,  $b_{ij}$ ) is a categorical variable that represents the longitudinal change of student  $i$  within school  $j$  who was at the top (bottom) of the distribution at grade 2.  $t_{ij}$  and  $b_{ij}$  are modelled as multinomial variables that take values in  $\{top, middle, bottom\}$  with probabilities  $\{\pi_{top,ij}, \pi_{middle,ij}, \pi_{bottom,ij}\}$  and  $\{\omega_{top,ij}, \omega_{middle,ij}, \omega_{bottom,ij}\}$ , respectively.

By considering the unvarying students as the reference category, the two multinomial multilevel models assume that:

$$\pi_{k,ij} = P(t_{ij} = k) = \frac{\exp(\varphi_{k,ij})}{1 + \sum_{k \in \{middle, bottom\}} \exp(\varphi_{k,ij})} \quad \text{for } k \in \{middle, bottom\} \quad (3)$$

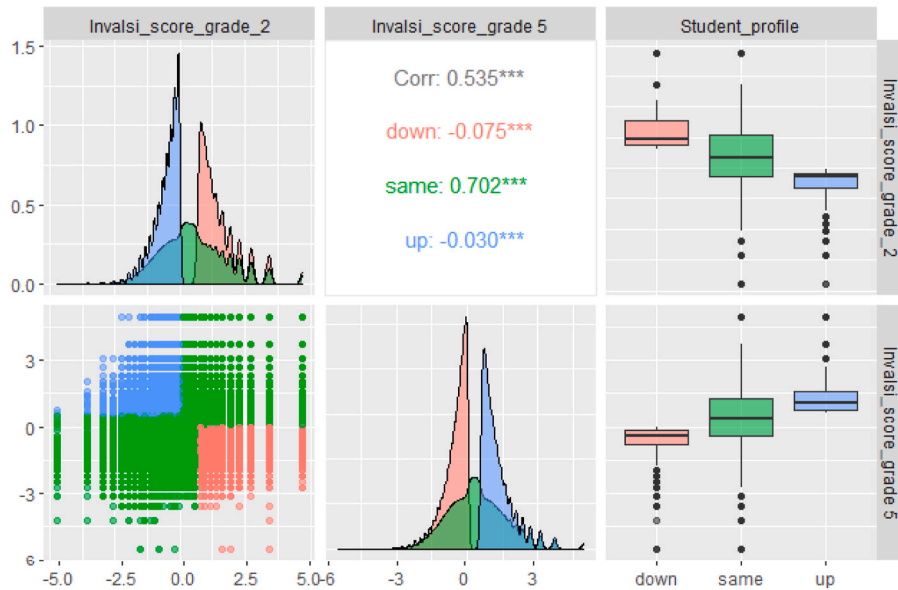
$$\omega_{h,ij} = P(b_{ij} = h) = \frac{\exp(\mu_{h,ij})}{1 + \sum_{h \in \{middle, top\}} \exp(\mu_{h,ij})} \quad \text{for } h \in \{middle, top\} \quad (4)$$

where  $\varphi_{k,ij} = \sum_{p=0}^p \beta_{pk} x_{ip} + s_{jk}$  and  $\mu_{h,ij} = \sum_{p=0}^p \alpha_{ph} x_{ip} + r_{jh}$  are the linear predictors.  $s_{jk} \sim N(0, \sigma_{sk}^2)$  and  $r_{jh} \sim N(0, \sigma_{rh}^2)$  are the random intercepts that represent the school effect on the probability of longitudinal changes of its students.

Both the probabilities to upgrade/downgrade are estimated by considering the following student covariates: students' personal characteristics (gender, socioeconomic status, nationality), previous education and academic performance (kindergarten attendance, early/late enrolment, average school score at grade 2, the delta, i.e., the variation, of the average school score between grades 2 and 5), class peers' information measured at grade 5 (peers' socioeconomic status, percentage of foreigners, percentage of females, number of students, delta, i.e., variation, in the number of students between grades 2 and 5, peers'

<sup>4</sup> The metric Area Under the Curve of the Receiver Operating Characteristic (AUC-ROC) is a widely used performance measure in machine learning and classification tasks. It quantifies the overall predictive power and discriminative ability of a model by evaluating the trade-off between true positive rate (sensitivity) and false positive rate (1 - specificity) across various classification thresholds. A higher AUC-ROC value, ranging from 0 to 1, indicates better model performance in distinguishing between different classes or categories.

(a) Mathematics



(b) Reading

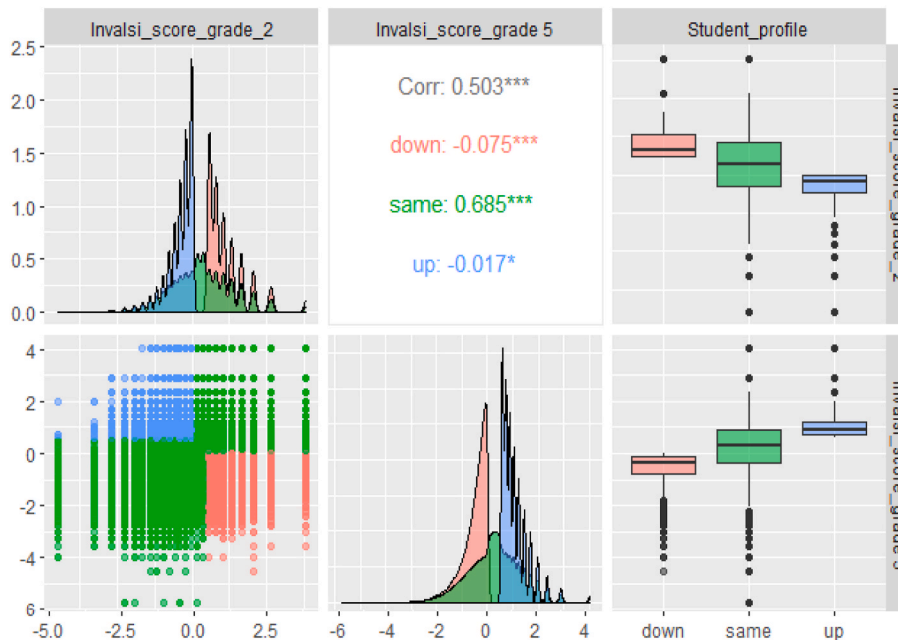


Fig. 1. INVALSI score distribution centred at 0 at grade 2 and 5, data at student level.

Note: The figure combines the names of the columns on the x-axis and the names of the rows on the y-axis, with colours indicating student profiles: blue refers to upgrading students, green to unvarying students, and pink to downgrading students. With this interpretation, the graph in the upper left corner shows the distribution of INVALSI scores centred at 0 at grade 2 by profile, with upgrading students at the bottom of the distribution when they were in grade 2, and downgrading students at the top. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

average score at grade 2, delta in peers' average score between grades 2 and 5) and environment information (geographical area, dummy for having changed school or region during primary school). The inclusion of many variables at this step is important for management and policy reasons.

While some of these variables are beyond the direct control of school principals and policymakers, they serve as essential controlling factors. Identifying the variables associated with changes in students'

persistence enables us to provide valuable insights for school leaders and policymakers. For instance, recognizing the impact of socioeconomic status or peers' performance can help in developing targeted interventions and support strategies.

Although we acknowledge that the correlations identified do not represent causal effects, they highlight areas where further investigation and data collection are needed. This includes variables such as the number and qualifications of teachers, which are not currently available

but could significantly enhance future analyses. In this vein, developing and collecting additional data and indicators is a key direction for future research, ultimately leading to more informed and effective educational policies.

#### 4. Results of the empirical analysis

##### 4.1. Longitudinal transitions and at-risk status

A considerable proportion of primary school students who struggled in mathematics or reading are at-risk of falling behind by grade 8. Specifically, 45 % of downgrading pupils in mathematics (equivalent to 8636 students) and 41 % of downgrading pupils in reading become at-risk students at grade 8 [Table 4]. This means that a negative transition in student performance during primary school significantly predicts the likelihood of struggling also during lower secondary school. On the other hand, a significant proportion of upgrading pupils (32 % in mathematics and 45 % in reading) perform above level 4 in grade 8. This is the first key message of our research.

We also consider the effects that major changes in performance during primary school can have on the performance at the end of lower secondary school. The model considering grade 5 covariates [Table 5] reveals that the higher a student’s test scores are at grade 5, the lower is this student’s likelihood of being at-risk at grade 8, especially when the student did not experience a longitudinal change during primary school, for both mathematics ( $\beta = -0.62$ ) and reading ( $\beta = -1.14$ ). For upgrading and downgrading students, having higher test scores at grade 5 is still a protective factor with a similar effect for downgrading in mathematics ( $\beta = -0.63$ ) and with a lower effect for all other cases ( $\beta = -0.05$  for upgrading in mathematics;  $\beta = -0.92$  and  $\beta = -0.031$  for downgrading and upgrading in reading).

This is the second key message of our research, and suggests that students who perform well during the entire primary school are more likely to perform better at grade 8 compared to their peers who have the same performance at grade 5 but experience up-and-down trends in time. These estimates change only slightly when updating the student’s academic results at grade 8, as observed in Table 6.

The effect of longitudinal transitions can be better observed in Fig. 2 that depicts the interaction between upgrading (blue line), downgrading (red line) and unvarying (grey line) students and their INVALSI score at grade 5 (x-axis) in relation to their probability of being at risk at grade 8 (y-axis). In reading, the most at-risk group includes students with consistently poor performance (unvarying, grey line), closely followed by those who experience a decline in performance (downgrading, red line). Similarly in mathematics, students with the highest probability of becoming at-risk are either those with consistently low performance or those experiencing a performance decline during primary school. Conversely, students with consistently high performance (unvarying with high scores at grade 5) exhibit the lowest risk. Performance upgrading has a more significant effect on reading than on mathematics, where always-top performers outdistance all other categories. In this perspective, school interventions that help students improve their performance early in their educational journey can have positive effects that persist over time and payoff later in terms of academic results.

When looking at the influence of personal characteristics, it is seen

**Table 4**

Absolute number of students for each transition category and proficiency level at grade 8 for each subject.

Transitions by subject in primary school and level in grade 8		Proficiency Level 1	Proficiency Level 2	Proficiency Level 3	Proficiency Level 4	Proficiency Level 5
Mathematics	Downgrading	3328	5308	6046	3309	1271
	Unvarying	28,238	50,506	63,785	50,382	48,201
	Upgrading	2760	4337	5409	3746	2139
Reading	Downgrading	2622	6417	8337	686	686
	Unvarying	20,064	47,459	73,732	63,048	31,798
	Upgrading	1157	3109	7340	7052	2183

Notes: Proficiency Level 1 and Level 2 refer to at-risk students.

**Table 5**

Results of the Step 1 logit model to predict at-risk students at grade 8 using grade 5 school covariates.

		Mathematics	Reading
	(Intercept)	7.24 ***	5.43 ***
ScoreG5: Longitudinal change(g5-g2)	ScoreG5: Unvarying	-0.62 ***	-1.14 ***
	ScoreG5: Downgrading	-0.63 ***	-0.92 ***
Personal characteristics	ScoreG5: Upgrading	-0.05 ***	-0.31 ***
	Gender (male)	-0.18 ***	0.27 ***
	Student SES - ESCS	-0.25 ***	-0.23 ***
	Nationality (1st gen. foreigner)	-0.14 **	-0.11 *
Instruction and previous academic performance	Nationality (2nd gen. foreigner)	-0.18 ***	-0.07 **
	TA marks – g5	-1.00 ***	-0.85 ***
	Delta TA marks (g2 – g5)	0.31 ***	0.26 ***
	Kindergarten attendance (yes)	0.03	0.06 **
	Enrollment (early)	-0.04	0.01
	Enrollment (late)	0.35 ***	0.19 **
	Class SES – ESCS	-0.17 ***	-0.18 ***
	Geo. macro-area (Islands)	1.18 ***	0.79 ***
	Geo. macro-area (South)	0.76 ***	0.51 ***
	Geo. macro-area (Northeast)	-0.39 ***	-0.21 ***
Environment	Geo. macro-area (Northwest)	-0.29 ***	-0.21 ***
	Changing school G2-G5 same region	-0.10	-0.05
	Changing region G2-G5	-0.02	-0.01

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘.’ 1.

that a higher socioeconomic status is consistently associated with a lower probability of being at risk across all settings [Table 5, Table 6], with a reduction in risk for students from more affluent classes as well. Generally, being male increases the probability of being at-risk in reading but decreases it in mathematics. When considering grade 8 covariates [Table 6], this gap further increases in mathematics and decreases in reading, in favor of male students. First and second-generation immigrant students, when controlling for all other factors, were found to have a slightly lower probability of being at risk in mathematics, which corroborates the idea that their primary disadvantage is in reading [50].

With respect to students’ previous education and achievements, the findings highlight that a student’s previous high achievements have a protective effect against the likelihood for him/her to become an at-risk student. However, students with variable evaluations during primary school are more likely to become at risk, which confirms the importance of consistent performance. Late school enrolment is identified as a risk factor, particularly for mathematics. It is crucial to note that late enrolment is often linked to other risk factors, such as health issues [51] or immigrant status [52]. Additionally, the effect of attending kindergarten on reading appears minimal and inconsistent, with no statistically significant impact when considering grade 8 covariates and only a marginal effect in the grade 5 model.

**Table 6**  
Results of the Step 1 logit model to predict at-risk students at grade 8 using grade 8 covariates.

		Mathematics		Reading	
	(Intercept)	8.00	***	6.99	***
ScoreG5: Longitudinal change(g5-g2)	ScoreG5: Unvarying	-0.56	***	-1.07	***
	ScoreG5:	-0.50	***	-0.81	***
Personal characteristics	Downgrading				
	ScoreG5: Upgrading	-0.01		-0.25	***
	Gender (male)	-0.38	***	0.05	***
	Student SES - ESCS	-0.11	***	-0.14	***
	Nationality (1st gen. foreigner)	-0.10	*	0.00	
	Nationality (2nd gen. foreigner)	-0.15	***	0.01	
Instruction and previous academic performance	TA marks - g8	-1.20	***	-1.12	***
	Delta TA marks (g5 - g8)	0.54	***	0.49	***
	Kindergarten attendance (yes)	-0.02		0.01	
	Enrollment (early)	0.04		0.08	
Environment	Enrollment (late)	0.35	***	0.23	***
	Class SES - ESCS	-0.50	***	-0.37	***
	Geo. macro-area (Islands)	1.20	***	0.82	***
	Geo. macro-area (South)	0.73	***	0.52	***
	Geo. macro-area (Northeast)	-0.40	***	-0.21	***
	Geo. macro-area (Northwest)	-0.33	***	-0.24	***
	Changing school G2-G5 same region	-0.21	**	-0.08	
	Changing region G2-G5	0.01		0.01	
	Changing school same region G5-G8	0.19	**	0.05	
	Changing region G5-G8	0.16	***	0.10	***

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

Regarding the geographical environment, there is huge territorial disparity. Students in the Italian islands (Sicily and Sardinia), with respect to those in the central regions, tend to be more likely to be at-risk in both subjects. Instead, in northern Italy students are less likely to be at-risk than their counterparts in central Italy. This finding underlines once again the North/South divide existing in the country's education system already highlighted in previous contributions [53–55]. Changing school or region in the transition between primary school and lower secondary school, also tends to increase the probability of being at risk in grade 8, especially in mathematics.

The analysis reveals that most of the covariates' effects are coherent between models at grade 5 and at grade 8, as shown in Tables 5 and 6. Moreover, the performance of the different models is also remarkably similar. Schools may already identify 77 % and 80 % of students who will be at risk in mathematics and reading at grade 8, respectively, based on the sole information available in grade 5, as shown in Table 7. Using grade 8 covariates only improves this capacity by 3 % for mathematics. This means that early models can accurately recognize students at-risk, providing a three-year window to carry out academic interventions. From a policy perspective, the timeliness of the prediction models is more crucial than their accuracy. Additionally, the performance obtained by early models is already substantial and comparable to that of later models. This is the third key message emerging from our study.

#### 4.2. Factors associated to longitudinal transitions at primary schools

The second step of the analysis highlights that the most significant factors influencing proficiency development in both mathematics and reading are students' previous achievements and students' peers' academic achievements over time [Table 8]. Two different models have

been tested, one comparing upgrading students to students whose performance was consistently at the bottom during primary school, and the other comparing downgrading students to consistently top-performing students. For the sake of brevity, Table 8 only shows the estimates for the logits relative to the biggest longitudinal changes, i.e., from top to bottom and vice-versa (the complete results are reported in Appendix C).

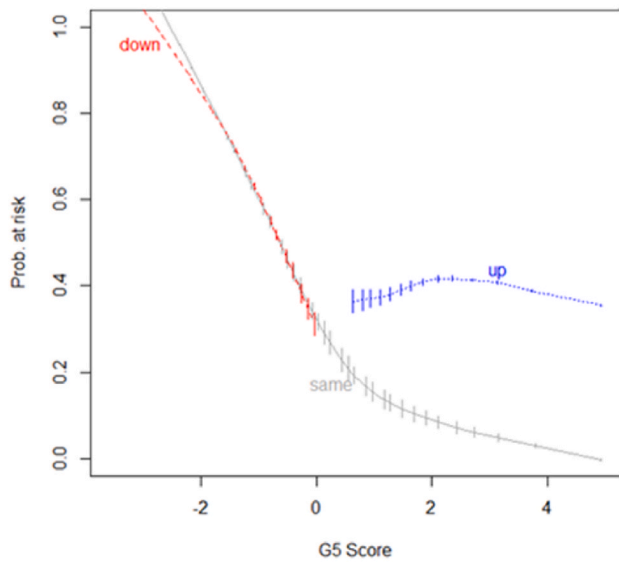
The higher the school performance at grade 2 of a bottom-student is, the higher is this student's probability of upgrading in both mathematics ( $\beta = 1.56$ ) and reading ( $\beta = 1.56$ ). On the other hand, the higher the school performance at grade 2 of a top-student is, the lower is this student's probability of downgrading in both mathematics ( $\beta = -1.96$ ) and reading ( $\beta = -1.85$ ). In this respect, attending a high-quality primary school may help students perform better in the early years of primary education which, according to these results, protects them from being at risk in lower secondary school. This factor might be crucial in discussions about school choice.

The peers' average score is also an important additional protective factor: the higher the peers' average score at grade 2 is, the lower is the probability to downgrade and the higher is the probability to upgrade. Moreover, the longitudinal transition of peers is associated with the longitudinal transition of individual students. A 1-point increase in the peers' scores between grades 2 and 5 results in an increase of 3.21-points for mathematics and 1.99-points for reading in the upgrade logit and a decrease of 2.85-points for mathematics and 1.82-points for reading in the downgrade logit. Thus, peer effect represents a relevant factor associated with the likelihood of a positive or a negative transition. This is particularly important to know in the context of the Italian educational system, where the peer group within a school is nested in classes that remain consistent between grade 1 and grade 5, and again between grade 6 and grade 8. These considerations should be kept in mind by principals, who are the main responsible actors for the class formation. A counterintuitive result is instead observed for the peers' SES that results to be a risk factor. This evidence could be induced by the peculiar subsample we are considering (only top/bottom performers at grade 2) and deserves further investigation.

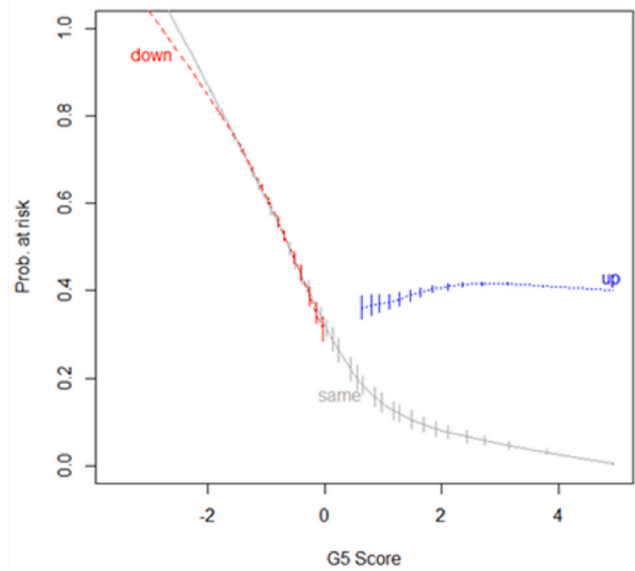
When considering students' personal characteristics, male students are more likely to improve their performance in mathematics ( $\beta = 0.51$ ) and less likely to decrease it ( $\beta = -0.71$ ), while no significant gender differences are observed in reading. This trend in mathematics is consistent with Contini et al. [56]. Additionally, the socioeconomic status plays a crucial role, as students with a higher socioeconomic background are more likely to show improvements, particularly in mathematics, and are less likely to experience a decline, especially in reading. In other words, socioeconomically advantaged students not only have better academic results, but they also improve them overtime at a faster pace. Moreover, second-generation immigrants are more likely to improve in mathematics, but also have a higher likelihood of experiencing a decline in reading performance, in line with the possible difficulty related to learning Italian as a second language.

The inclusion of a random intercept at the school level allows us to quantify the extent to which unexplained variation in students' transitions can be attributed to differences between schools. This is measured by the Percentage of Variability Explained by the Random Effects (PVRE). We observed a substantial level of heterogeneity across schools, particularly in mathematics, where the PVRE was found to be 15 % for students transitioning to higher levels and 20 % for those transitioning to lower levels. In reading, the variability explained by schools was 11 % for students upgrading and 15 % for students downgrading. These findings suggest that schools play a significant role not only through the influence of individual classes, as evidenced by the peers' covariates mentioned earlier, but also through other factors which in this research remain unobservable also because of a lack of school-related variables. Furthermore, upon analyzing the correlation of schools' random intercepts (see Fig. 3), an intriguing pattern emerges. Notably, schools that show a positive impact in one discipline tend to show a positive impact in the other discipline as well. Conversely, schools that have a negative

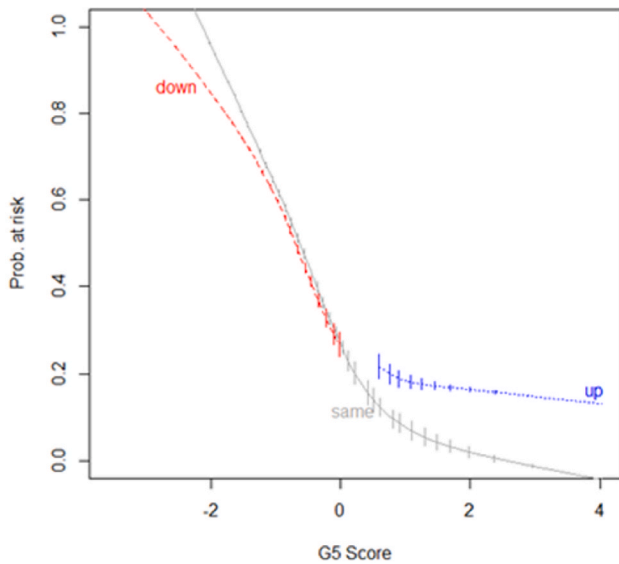
(a) Math – G5 covariates



(b) Math – G8 covariates



(c) Reading – G5 covariates



(d) Reading – G8 covariates

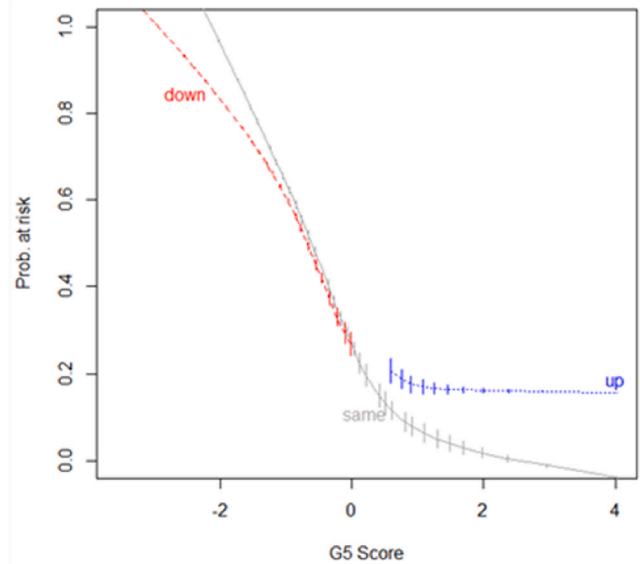


Fig. 2. Interaction between the score at grade 5 and the longitudinal transitions during primary school (up, down or unvarying) and the likelihood of becoming at-risk (grade 8).

impact in one discipline also tend to have a negative impact in the other discipline. This consistent pattern across disciplines strongly suggests that schools' influence extends beyond individual disciplines, indicating a broad effect on students' academic performance. In line with the literature framing at-risk students as a systemic failure (V. E. [47]), it is evident that schools have significant effects on students' transitions, beyond individual behaviors and backgrounds. Given our access to a limited number of variables describing school characteristics, a promising avenue for future research involves expanding the range and types of variables to further explore the nature and mechanisms of schools' effects.

Finally, we are interested in characterizing students who exhibit longitudinal transitions in both the subjects, distinguishing between those who improve from those who decline. We selected two distinct

samples: students in the bottom 40 % of the distribution in both mathematics and reading at grade 2 (71,819 students, 25.76 % of the total sample) and students who were in the top 40 % (78,341 students, 28.10 % of the total sample). For each of these groups, we model the probability of transitioning either to the top 40 % or the bottom 40 % distribution in both subjects by grade 5. The bottom grade 2 students who reach the top 40 % in both the subjects at grade 5 are 3912 (5.4 %), while the top students who transit to the bottom 40 % are 4951 (6.3 %). We fit two logistic regression models considering all variables used in the MCMCgmm model as covariates. For this specific analysis, due to the low number of observations within schools, we do not take into account of the nested structure of students within schools. Results of the reduced models obtained by the backward stepwise regressions are reported in Table 9. We observe that the individual and the peers average

**Table 7**  
Step 1 logit models performances computed on the test dataset.

Model metrics	Mathematics		Reading	
	Grade 5	Grade 8	Grade 5	Grade 8
Accuracy	76.4 %	78.0 %	76.9 %	78.3 %
Sensitivity	77.1 %	80.5 %	80.1 %	80.7 %
Specificity	86.6 %	88.4 %	90.2 %	90.6 %
Precision	62.7 %	64.3 %	65.8 %	64.7 %
AUC-ROC	0.84	0.87	0.86	0.87
Cut-off	0.33	0.33	0.26	0.27

Notes: Accuracy: (TP + TN)/(TP + TN + FP + FN); Sensitivity: TP/(TP + FN); Specificity: TN/(TN + FP); Precision: TP/(TP + FP). AUC-ROC: trade-off between true positive rate (sensitivity) and false positive rate (1 - specificity) across various classification thresholds; Cut-off: optimal threshold.

TP: true positives; TN: true negatives; FP: false positives; FN: False negatives.

school scores at grade 2 are relevant and protective factors (the higher (lower) the student and peers scores, the higher (lower) the probability to upgrade (downgrade)). In particular, the g5-g2 delta of class scores is positively associated with the type of student transition, indicating that the student is likely to upgrade/downgrade together with the rest of the class. Changing the school or the region during primary school increases the probability of having a longitudinal transition (of any type), suggesting that the student is less likely to have a stable performance. In the MCMC setting, the changing school coefficients may have been captured by the measured school-effect. The percentages of immigrant students and female students is found to be significant only in their association with the student upgrade probability, with a positive effect for the percentage of immigrants and a negative effect for the percentage of females. Conversely, gender and being a late/early enrolled student only play a role in the downgrade likelihood, where being a male student and an early enrolled student are protective factors, while late enrolment is a risk factor. The student SES is positively (negatively) associated to the upgrade (downgrade) probability, while the class SES has an opposite effect. As mentioned earlier, this effect sounds counterintuitive and deserves further investigation. Lastly, geographical areas are also relevant in longitudinal transitions. Being a student in the Islands or in

**Table 8**  
Step 2 multilevel model to explore the factors associated with proficiency level transitions in Mathematics and Reading between grade 5 and 8.

Subject	Mathematics		Reading	
	Upgrading	Downgrading	Upgrading	Downgrading
Longitudinal changes				
(Intercept)	-15.12 ***	13.12 ***	-13.8 ***	12.68 ***
Personal characteristics				
Gender (male)	0.51 ***	-0.71 ***	0.04 .	0.01
Student SES	0.13 ***	-0.24 ***	0.21 ***	-0.28 ***
Nationality (1st gen. foreigner)	0.12	-0.25 .	0.07 **	0.01
Nationality (2nd gen. foreigner)	0.23 ***	-0.06	-0.11	0.16 **
Instruction and previous academic performance				
TA marks - grade 2	1.56 ***	-1.96 ***	1.56 ***	-1.85 ***
Delta TA marks (g2-g5)	1.19 ***	-1.41 ***	1.14 ***	-1.31 ***
Kindergarten attendance (yes)	0.08	-0.08	0.05	-0.11 *
Enrollment (early)	0.15	-0.24 *	-0.01	-0.04
Enrollment (late)	-0.18	0.29 .	-0.23 .	0.17
Peers' characteristics (grade 5) and peers' previous performances				
Peers SES	-0.47 ***	0.44 ***	-0.38 ***	0.38 ***
% peers 1st gen. foreigner	0.01 .	-0.01 .	0.01 **	-0.01 *
% peers 2nd gen. foreigner	0	0	0.01 ***	-0.01 ***
% peers female	0	0	-0.01 ***	0 ***
No. students	0	0.02 ***	-0.01 *	0.01 ***
Delta No. students (g2-g5)	0	0	0	0
Peer score - grade 2	1.54 ***	-2.51 ***	1.06 ***	-1.6 ***
Delta peer score (g2-g5)	3.21 ***	-2.85 ***	1.99 ***	-1.82 ***
Environment				
Geo. macro-area (Islands)	0.14 *	0.38 ***	-0.03	0.46 ***
Geo. macro-area (South)	0.4 ***	-0.05	0.21 ***	0.02
Geo. macro-area (Northeast)	-0.49 ***	0.04	-0.29 ***	0.04
Geo. macro-area (Northwest)	-0.48 ***	0.07	-0.34 ***	0.08 .
Changing school g2-g5	0.07 .	0.04	0.08 *	0.03
Changing region g2-g5	0.29 .	-0.036	0.29 .	-0.04

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

southern Italy increases the probability to have a longitudinal transition (of any type), reflecting greater heterogeneity in the educational process in these geographical areas, while students in northern Italy are on average more inclined to have a stable performance across the different grades.

Overall, these results are coherent with the marginal ones obtained using the MCMCglmm model, reported in Table 8. However, they also uncover specific nuances in predictors when examining transitions in both subjects collectively. Notably, there is a reduction in gender differences and an absence of impact for immigrant students. Simultaneously, there is a more consistent influence of environmental transitions possibly explained by the dissipation of school effects into broader environmental factors. Similar to the MCMCglmm model, the persistence of peers' transitions over time emerges as the most robust predictor. In the joint context, i.e., when modelling the transition in both subjects, the effects vary depending on the type of transition. For upgrading students, an improvement in peers' reading ability has a stronger impact, increasing the likelihood of student improvement. Conversely, an improvement in peers' math ability acts as a stronger protective factor against downgrading. This highlights the critical importance of delving into subject-specific and combined models to gain a comprehensive understanding of the multifaceted factors shaping students' academic trajectories.

### 5. Discussion and conclusive remarks

Tracking student progress and evaluating achievement growth over time play a critical role in the instructional feedback loop, offering valuable evidence on academic development that can be utilized to assess and refine existing policies and practices [38]. This study highlights the significance of monitoring students' performance and proficiency transitions during their primary school years as a means to identify students at-risk at the end of lower secondary school. Proficiency transitions in primary school emerge as a significant predictor for identifying at-risk students by grade 8, indicating that early identification of these students in the Italian context can be achieved as early as grade 5. With this evidence, we do not want to promote a

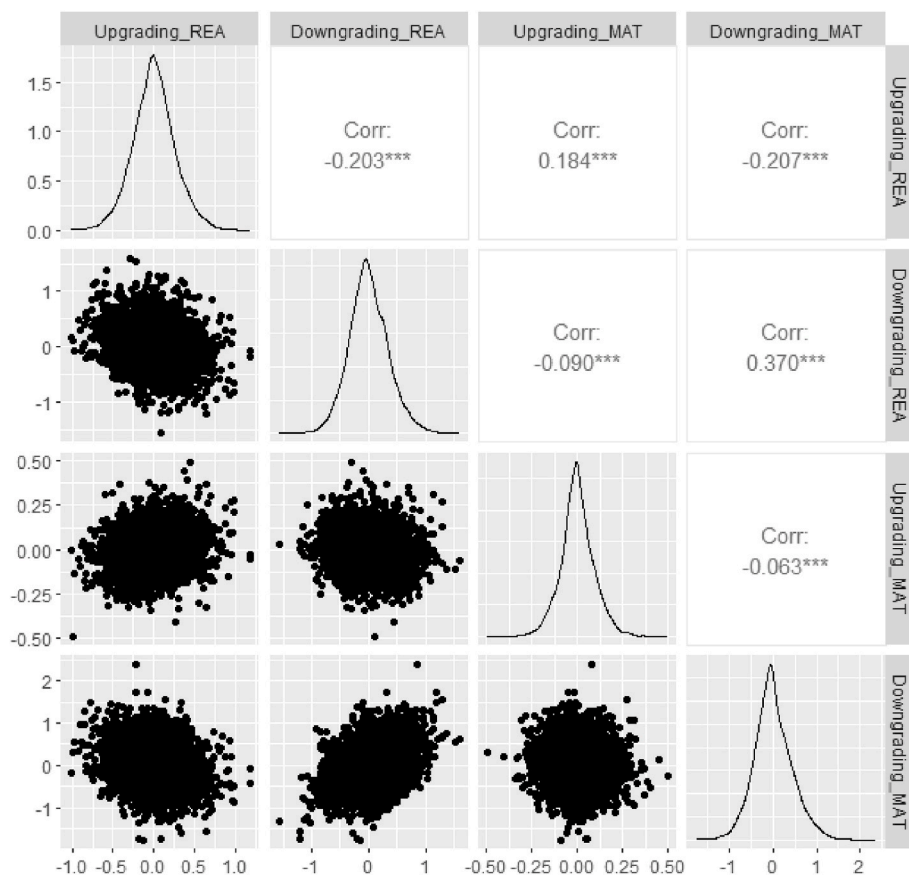


Fig. 3. Distributions of school effects on the probability of student transitions, estimated by the step 2 multilevel model, for reading and mathematics.

“deterministic” view of the educational path, where early stages are necessarily conducive to subsequent good or bad performance. Instead, we want to provide empirical evidence about the “cumulative” nature of education, acknowledging the importance of supporting students with specific needs as soon as possible. In other words, our research claims not for tracking students, but instead for offering equal opportunities to all. It is then school responsibility to identify the best (personalized) interventions for the students who are identified as more at-risk (by the way, the same reasoning holds for high-performing students). In this vein, the method proposed here is an instrument for supporting school awareness and autonomy in promoting and realizing individualized educational interventions and programs.

Through an analysis of factors associated with students’ transitions in primary school utilizing the Markov Chain Monte Carlo (MCMC) glmm method, we see that predictors for upgrading and downgrading performance share similarities, but their magnitudes differ significantly. In particular, peers exert a more marked influence on upgrading students’ performance compared to downgrading them. A one-standard deviation increase in peers’ performance substantially enhances the likelihood of a student to improve his/her own performance, while also providing a moderate protective effect against downgrading. This is in line with previous evidence on peer effects in similar contexts [57]. This strong class-level factor is potentially related, to some extent, to the quality of teachers, a dimension not available when using data provided by INVALSI or by other Italian educational entities at a national scale.

Additionally, the institutional impact on proficiency development is evident from the variability explained by heterogeneity across schools. Schools that demonstrate a positive/negative impact in one subject tend to exhibit a similar effect in the other subject, indicating an influence on students’ academic performance that goes beyond the individual subject. This finding also suggests that schools are already playing a substantial role in impacting the educational performance and trajectories

of their students. Any intervention to be developed following the identification of at-risk students should move from this awareness.

According to our findings, potential interventions to prevent students downgrading and/or to support student upgrading during primary school can be applied at both class and school levels. Students who belong to downgrading classes and/or schools which have an overall negative effect on pupils’ performance can largely benefit from early educational interventions and continuous monitoring of proficiency changes. Early identification and continuous monitoring of longitudinal changes in students’ proficiency have a great potential for mitigating disadvantages in later stages of the educational path and enabling social mobility. In line with our results, changing students across classes during primary school might be a strategy to be considered for supporting at-risk students. As a limitation, we acknowledge the fact that, despite the data collection being censorial, technical issues in tracking the student identifier bring a loss of observation that might be not completely at random. Moreover, despite being unlikely, we cannot completely rule out the possibility of some mechanisms of self-selection of students into high performing classes being at play in lower secondary schools, thus influencing the peer composition of classes. Additionally, while our study finds that peers whose performance declines increase the likelihood of students being non-proficient on standardized tests, it is important to note that peer effects related to declining performance can have various implications. For instance, a student’s performance may appear relatively better in a low-performing group and worse in a high-performing group. These dynamics could influence grade repetition, dropout rates, and teacher-assigned grades differently from standardized test scores.

This study emphasizes the importance of understanding the key factors that contribute to proficiency development. By identifying at-risk students and implementing targeted interventions at the class and school levels, policymakers and educators can refine existing policies

**Table 9**  
Results of the step 2 auxiliary logit models to explore the factors associated with proficiency level transitions in both subjects between grade 5 and 8.

	<i>Transition</i>	<i>Upgrading</i>	<i>Downgrading</i>		
	(Intercept)	−0.198 ***	0.725 ***		
Personal characteristics	Gender (male)		−0.005 ***		
	Student SES - ESCS	0.008 ***	−0.012 ***		
Instruction and previous academic performance	TA marks grade 2	0.025 ***	−0.050 ***		
	Delta TA marks Mathematics	0.020 ***	−0.036 ***		
	TA marks Reading – grade 2	0.010 ***	−0.027 ***		
	Delta TA marks Reading	0.013 ***	−0.022 ***		
	Enrollment (early)		−0.015 **		
	Enrollment (late)		0.032 ***		
	Peers' characteristics (grade 5) and peers' previous performances	Peers SES	−0.047 ***	0.023 ***	
		% peers 1st gen. foreigner	0.072 ***		
		% peers 2nd gen. foreigner	0.045 ***		
		% peers female	−0.061 ***		
Delta No. students (g2-g5)			0.008 **		
Peer score Mathematics - grade 2		0.045 ***	−0.036 ***		
Delta peer score Mathematics (g2-g5)		0.080 ***	−0.065 ***		
Peer score Reading - grade 2		0.078 ***	−0.017 ***		
Delta peer score Reading (g2-g5)		0.114 ***	−0.045 ***		
Environment		Geo. macro-area (Islands)	0.039 ***	0.038 ***	
	Geo. macro-area (South)	0.026 ***	0.016 ***		
	Geo. macro-area (Northeast)	−0.015 ***	−0.012 ***		
	Geo. macro-area (Northwest)	−0.012 ***	−0.006 **		
	Changing school g2-g5	0.005 **	0.008 ***		
	Changing region g2-g5	0.041 ***	0.030 ***		

Area Under Curve (AUC) 0.901 0.905.

Observations 71,819 78,341

and practices to ensure better academic outcomes for all students, taking

**Appendix A. Longitudinal transitions in primary school and risk identification**

The threshold we use to identify longitudinal transitions in primary school corresponds to the level of improvement or decline in primary school students' performance that best predicts students' level of risk at grade 8. To accomplish this, we compared the effectiveness of binary logit models that consider various longitudinal transitions in primary school to predict the probability of being at risk at grade 8. The optimal threshold, for both mathematics and reading, results to be 40 %, identifying students who move from the top 40 % at grade 2 to the bottom 40 % at grade 5 and vice versa [Fig. 1].

timely action.

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

**CRedit authorship contribution statement**

**Melisa L. Diaz Lema:** Writing – review & editing, Writing – original draft, Validation, Investigation, Formal analysis, Data curation, Conceptualization. **Chiara Masci:** Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation. **Mara Soncin:** Writing – review & editing, Validation, Formal analysis, Data curation, Conceptualization. **Tommaso Agasisti:** Writing – review & editing, Validation, Supervision.

**Acknowledgments**

We thank the anonymous reviewers for their thoughtful comments, as well as SEPS Editor-in-Chief Rajan Batta and the SEPS support staff for their valuable assistance. We are also grateful to the participants of the Lisbon Economics and Statistics of Education conferences and the seminar “INVALSI Data: A Tool for Teaching and Scientific Research” for their insightful feedback. Chiara Masci acknowledges financial support from the Italian Ministry of University and Research (MUR) under the Department of Excellence 2023-2027 grant agreement “Centre of Excellence in Economics and Data Science” (CEEDS).

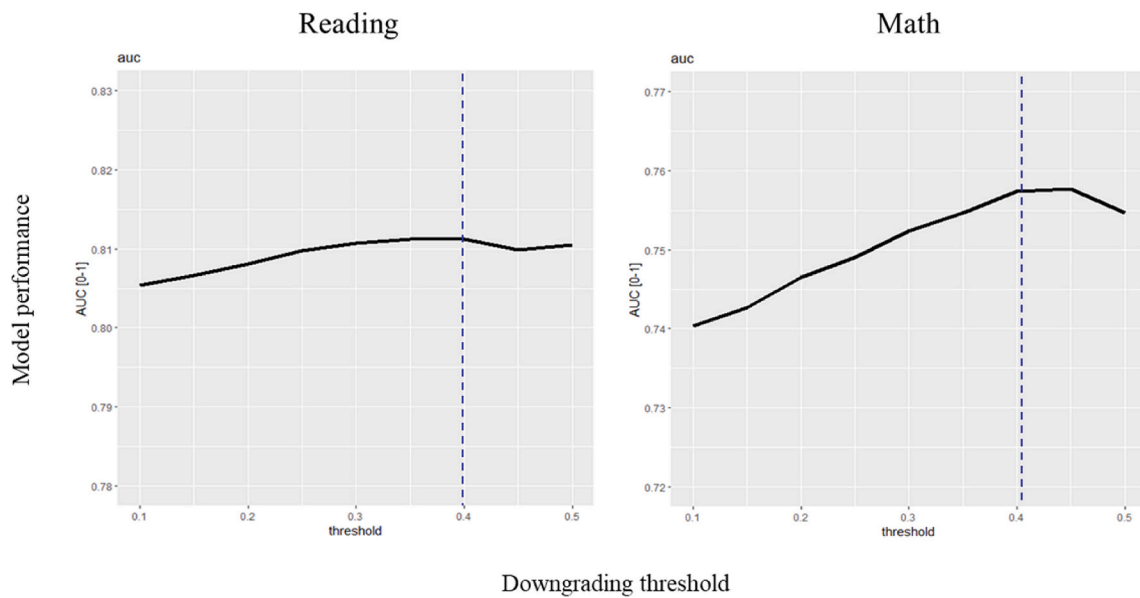


Fig. 1. Relationship between longitudinal transitions thresholds (represented on the x-axis) and model performance (measured on the y-axis with AUC-ROC<sup>5</sup>) for identifying at-risk students in both reading and mathematics.

**Appendix B. Distribution of study sample characteristics, including means and standard deviations for numerical variables and proportions for categorical variables, compared to the entire population**

Category	Variable	Type	Full data, N = 403,786 <sup>a</sup>	Study sample, N = 278,765 <sup>a</sup>
Outcome <sup>a</sup>	at-risk math	Binary	0.34 (0.47)	0.34 (0.47)
	at-risk read.	Binary	0.29 (0.46)	0.29 (0.46)
Personal characteristics	gender	Cat	50 %	50 %
	nationality	Cat		
	-Italian		92 %	93 %
	-1 <sup>st</sup> gen. immigrant		1.7 %	1.6 %
	-2 <sup>nd</sup> gen. immigrant		6.0 %	5.8 %
Instruction and previous academic performance	Socio-economic status (ESCS index)	Num	0.11 (0.92)	0.12 (0.92)
	kindergarten attendance	Cat	90 %	91 %
	enrollment	Cat		
	-regular		98 %	98 %
	-early		1.3 %	1.3 %
	-late		1.0 %	0.9 %
	TA marks math G2	Num	8.16 (1.02)	8.19 (1.00)
	TA marks read. G2	Num	8.05 (1.03)	8.08 (1.02)
	Test score math G2	Num	0.49 (1.25)	0.51 (1.26)
	Test score read. G2	Num	0.31 (1.01)	0.33 (1.01)
	TA marks math G5	Num	7.98 (1.07)	8.00 (1.06)
	TA marks read. G5	Num	7.92 (1.01)	7.94 (1.01)
	Test score math G5	Num	0.40 (1.20)	0.41 (1.20)
Test score read. G5	Num	0.25 (1.01)	0.27 (1.01)	
TA marks math G8	Num	6.90 (1.39)	6.93 (1.39)	
TA marks read. G8	Num	7.09 (1.16)	7.11 (1.16)	
Environment	reg. location	Cat		
	-center		18 %	16 %
	-islands		9.9 %	10 %
	-northeast		19 %	20 %
	-northwest		27 %	27 %
	-south		26 %	27 %

<sup>a</sup> Cat: %; Numeric: Mean (SD).

**Appendix C. Results of the step 2 multilevel model to explore the factors associated with proficiency level transitions in a) Mathematics and b) Reading between grade 5 and 8**

a) Mathematics									
Longitudinal changes		Downgrading		Drop to sufficiency		Reach sufficiency		Upgrading	
	(Intercept)	13.12	***	7	***	-9.82	***	-15.12	***
Personal characteristics	Gender (male)	-0.71	***	-0.43	***	0.36	***	0.51	***

(continued on next page)

(continued)

a)Mathematics		Downgrading		Drop to sufficiency		Reach sufficiency		Upgrading	
Longitudinal changes									
(Intercept)		13.12	***	7	***	-9.82	***	-15.12	***
Instruction and previous academic performance	Student SES	-0.24	***	-0.14	***	0.06	***	0.13	***
	Nationality (1st gen. foreigner)	-0.25	.	-0.23	.	0.03	.	0.12	.
	Nationality (2nd gen. foreigner)	-0.06	.	-0.09	.	0.05	.	0.23	***
	TA marks - grade 2	-1.96	***	-1.08	***	1.01	***	1.56	***
	Delta TA marks (g2-g5)	-1.41	***	-0.82	***	0.73	***	1.19	***
	Kindergarten attendance (yes)	-0.08	.	-0.01	.	0.09	<sup>a</sup>	0.08	.
Peers' characteristics (grade 5) and peers' previous performances	Enrollment (early)	-0.24	<sup>a</sup>	-0.1	.	0.27	<sup>a</sup>	0.15	.
	Enrollment (late)	0.29	.	0.22	.	-0.23	.	-0.18	.
	Peers SES	0.44	***	0.24	***	-0.25	***	-0.47	***
	% peers 1st gen. foreigner	-0.01	.	0	.	0	.	0.01	.
	% peers 2nd gen. foreigner	0	.	0	<sup>a</sup>	0	.	0	.
	% peers female	0	.	0	.	0	<sup>a</sup>	0	.
	No. students	0.02	***	0.01	***	0	.	0	.
	Delta No. students (g2-g5)	0	.	0	.	0	.	0	<sup>a</sup>
	Peer score - grade 2	-2.51	***	-1.45	***	0.82	***	1.54	***
	Delta peer score (g2-g5)	-2.85	***	-1.61	***	1.74	***	3.21	***
Environment	Geo. macro-area (Islands)	0.38	***	0.19	***	0.06	.	0.14	<sup>a</sup>
	Geo. macro-area (South)	-0.05	.	-0.07	.	0.25	***	0.4	***
	Geo. macro-area (Northeast)	0.04	.	0.06	.	-0.17	**	-0.49	***
	Geo. macro-area (Northwest)	0.07	.	0.12	**	-0.2	***	-0.48	***
	Changing school g2-g5	0.04	.	-0.02	.	0.06	.	0.07	.
	Changing region g2-g5	-0.04	.	0.1	.	-0.13	.	0.29	.

b)Reading		Downgrading		Drop to sufficiency		Reach sufficiency		Upgrading	
Longitudinal changes									
(Intercept)		12.68	***	6.62	***	-8.85	***	-13.8	***
Personal characteristics	Gender (male)	0.01	.	0.02	.	0	.	0.04	.
	Student SES	-0.28	***	-0.19	***	0.12	***	0.21	***
	Nationality (1st gen. foreigner)	0.01	.	-0.05	.	-0.06	.	0.07	**
Instruction and previous academic performance	Nationality (2nd gen. foreigner)	0.16	**	0.04	.	-0.07	.	-0.11	.
	TA marks - grade 2	-1.85	***	-1.02	***	1	***	1.56	***
	Delta TA marks (g2-g5)	-1.31	***	-0.72	***	0.71	***	1.14	***
	Kindergarten attendance (yes)	-0.11	<sup>a</sup>	0	.	0.07	.	0.05	.
	Enrollment (early)	-0.04	.	0.04	.	0.09	.	-0.01	.
	Enrollment (late)	0.17	.	-0.01	.	-0.12	.	-0.23	.
Peers' characteristics (grade 5) and peers' previous performances	Peers SES	0.38	***	0.23	***	-0.23	***	-0.38	***
	% peers 1st gen. foreigner	-0.01	<sup>a</sup>	-0.01	**	0	.	0.01	**
	% peers 2nd gen. foreigner	-0.01	***	0	.	0	.	0.01	***
	% peers female	0	***	0	***	0	***	-0.01	***
	No. students	0.01	***	0.01	***	-0.01	**	-0.01	<sup>a</sup>
	Delta No. students (g2-g5)	0	.	0	.	0	.	0	.
	Peer score - grade 2	-1.6	***	-1	***	0.55	***	1.06	***
Environment	Delta peer score (g2-g5)	-1.82	***	-1.05	***	0.97	***	1.99	***
	Geo. macro-area (Islands)	0.46	***	0.22	***	-0.2	***	-0.03	.
	Geo. macro-area (South)	0.02	.	-0.01	.	0.03	.	0.21	***
	Geo. macro-area (Northeast)	0.04	.	0.11	<sup>a</sup>	-0.13	**	-0.29	***
	Geo. macro-area (Northwest)	0.08	.	0.13	***	-0.14	***	-0.34	***
	Changing school g2-g5	0.03	.	-0.02	.	0.03	.	0.08	<sup>a</sup>
	Changing region g2-g5	-0.04	.	-0.2	.	0.38	**	0.29	.

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

**Data availability**

The authors do not have permission to share data.

**References**

[1] Yi HS, Na W, Lee C. Predicting secondary school students' academic achievement from their elementary school performance and learning behaviours: a longitudinal study based on South Korea's national assessment of educational achievement. *Asia Pac J Educ* 2021;1-18. <https://doi.org/10.1080/02188791.2021.1963673>.

[2] OECD. Equity and quality in education: supporting disadvantaged students and schools. OECD; 2012. <https://doi.org/10.1787/9789264130852-en>.

[3] Glaeser J. Relative educational poverty: conceptual and empirical issues. *Qual Quantity* 2022;56(4):2803-20. <https://doi.org/10.1007/s11135-021-01226-3>.

[4] Cunha F, Heckman J. The technology of skill formation. w12840. National Bureau of Economic Research; 2007, w12840. <https://doi.org/10.3386/w12840>.

[5] Freeman J, Simonsen B. Examining the impact of policy and practice interventions on high school dropout and school completion rates: a systematic review of the literature. *Rev Educ Res* 2015;85(2):205-48. <https://doi.org/10.3102/0034654314554431>.

[6] Mac Iver MA, Stein ML, Davis MH, Balfanz RW, Fox JH. An efficacy study of a ninth-grade early warning indicator intervention. *J Res Educational Effect* 2019;12(3):363-90. <https://doi.org/10.1080/19345747.2019.1615156>.

[7] Sansone D. Beyond early warning indicators: high school dropout and machine learning. *Oxf Bull Econ Stat* 2019;81(2):456-85. <https://doi.org/10.1111/obes.12277>.

[8] Davis M, Herzog L, Legters N. Organizing schools to address early warning indicators (EWIs): common practices and challenges. *J Educ Stud Placed A T Risk* 2013;18(1):84-100. <https://doi.org/10.1080/10824669.2013.745210>.

[9] Betebenner D. Norm- and criterion-referenced student growth. *Educ Meas* 2009;28(4):42-51. <https://doi.org/10.1111/j.1745-3992.2009.00161.x>.

[10] Doyle O, Harmon CP, Heckman JJ, Tremblay RE. Investing in early human development: timing and economic efficiency. *Econ Hum Biol* 2009;7(1):1-6. <https://doi.org/10.1016/j.ehb.2009.01.002>.

- [11] Xenidou-Dervou I, Van Luit JEH, Kroesbergen EH, Friso-van den Bos I, Jonkman LM, van der Schoot M, van Lieshout ECDM. Cognitive predictors of children's development in mathematics achievement: a latent growth modeling approach. *Dev Sci* 2018;21(6):e12671. <https://doi.org/10.1111/desc.12671>.
- [12] Anderman EM, Gimbirt B, O'Connell AA, Riegel L. Approaches to academic growth assessment. *Br J Educ Psychol* 2015;85(2):138–53. <https://doi.org/10.1111/bjep.12053>.
- [13] Aitkin M, Longford N. Statistical modelling issues in school effectiveness studies. *J Roy Stat Soc* 1986;149(1). <https://doi.org/10.2307/2981882>.
- [14] OECD. PISA 2012 results: ready to learn (volume III): students' engagement, drive and self-beliefs. OECD; 2013. <https://doi.org/10.1787/9789264201170-en>.
- [15] Turner R. *Described proficiency scales and learning metrics* (4; Assessment GEMs). Australian Council for Educational Research; 2014. <https://research.acer.edu.au/assessgems/4/>.
- [16] Frogner L, Hellfeldt K, Ångström A-K, Andershed A-K, Källström Å, Fanti KA, Andershed H. Stability and change in early social skills development in relation to early school performance: a longitudinal study of A Swedish cohort. *Early Educ Dev* 2022;33(1):17–37. <https://doi.org/10.1080/10409289.2020.1857989>.
- [17] Romano E, Babchishin L, Pagani LS, Kohen D. School readiness and later achievement: replication and extension using a nationwide Canadian survey. *Dev Psychol* 2010;46(5):995–1007. <https://doi.org/10.1037/a0018880>.
- [18] Klimstra TA, Luyckx K, Germeijs V, Meeus WHJ, Goossens L. Personality traits and educational identity formation in late adolescents: longitudinal associations and academic progress. *J Youth Adolesc* 2012;41(3):346–61. <https://doi.org/10.1007/s10964-011-9734-7>.
- [19] Masci C, De Witte K, Agasisti T. The influence of school size, principal characteristics and school management practices on educational performance: an efficiency analysis of Italian students attending middle schools. *Soc Econ Plann Sci* 2018;61:52–69. <https://doi.org/10.1016/j.seps.2016.09.009>.
- [20] Quinnan TW. *Adult students at-risk: culture bias in higher education*. Greenwood Publishing Group; 1997.
- [21] Vinas-Forcade J, Mels C, Van Houtte M, Valcke M, Derluyn I. Can failure be prevented? Using longitudinal data to identify at-risk students upon entering secondary school. *Br Educ Res J* 2021;47(1):205–25. <https://doi.org/10.1002/berj.3683>.
- [22] Gibbs BG, Heaton TB. Drop out from primary to secondary school in Mexico: a life course perspective. *Int J Educ Dev* 2014;36:63–71. <https://doi.org/10.1016/j.ijedudev.2013.11.005>.
- [23] Borgna C, Struffolino E. Pushed or pulled? Girls and boys facing early school leaving risk in Italy. *Soc Sci Res* 2017;61:298–313. <https://doi.org/10.1016/j.ssresearch.2016.06.021>.
- [24] Lee J-W, Barro RJ. Schooling quality in a cross-section of countries. *Economica* 2001;68(272):465–88. <https://doi.org/10.1111/1468-0335.00257>.
- [25] Rumberger RW. Dropping out of middle school: a multilevel analysis of students and schools. *Am Educ Res J* 1995;32(3):583–625. <https://doi.org/10.3102/00028312032003583>.
- [26] Mahuteau S, Mavromaras K. An analysis of the impact of socio-economic disadvantage and school quality on the probability of school dropout. *Educ Econ* 2014;22(4):389–411. <https://doi.org/10.1080/09645292.2014.918586>.
- [27] Aucejo EM, Romano TF. Assessing the effect of school days and absences on test score performance. *Econ Educ Rev* 2016;55:70–87. <https://doi.org/10.1016/j.econedurev.2016.08.007>.
- [28] Wöbmann L. Schooling resources, educational institutions and student performance: the international evidence. *Oxf Bull Econ Stat* 2003;65(2):117–70. <https://doi.org/10.1111/1468-0084.00045>.
- [29] Clotfelter CT, Ladd HF, Vigdor JL. Teacher credentials and student achievement: longitudinal analysis with student fixed effects. *Econ Educ Rev* 2007;26(6):673–82. <https://doi.org/10.1016/j.econedurev.2007.10.002>.
- [30] Löwe M, Rinne U, Sonnabend H. Gender role models and early-career decisions. *Appl Econ Lett* 2023;30(11):1526–30. <https://doi.org/10.1080/13504851.2022.2066618>.
- [31] Hill DV, Hughes RP, Lenard MA, Liebowitz DD, Page LC. New schools and new classmates: the disruption and peer group effects of school reassignment. *Econ Educ Rev* 2023;92:102316. <https://doi.org/10.1016/j.econedurev.2022.102316>.
- [32] Neild RC, Stoner-Eby S, Furstenberg F. Connecting entrance and departure: the transition to ninth grade and high school dropout. *Educ Urban Soc* 2008;40(5):543–69. <https://doi.org/10.1177/0013124508316438>.
- [33] Ehler M, Koedel C, Parsons E, Podgursky M. Selecting growth measures for use in school evaluation systems: should proportionality matter? *Educ Policy* 2016;30(3):465–500. <https://doi.org/10.1177/0895904814557593>.
- [34] Hamre BK, Pianta RC. Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Dev* 2001;72(2):625–38.
- [35] Albert WD, Hanson JL, Skinner AT, Dodge KA, Steinberg L, Deater-Deckard K, Bornstein MH, Lansford JE. Individual differences in executive function partially explain the socioeconomic gradient in middle-school academic achievement. *Dev Sci* 2020;23(5). <https://doi.org/10.1111/desc.12937>.
- [36] Goldstein H. Some models for analysing longitudinal data on educational attainment. *J Roy Stat Soc* 1979;142(4):407–42. <https://doi.org/10.2307/2982551>.
- [37] Mok MMC, McInerney DM, Zhu J, Or A. Growth trajectories of mathematics achievement: longitudinal tracking of student academic progress. *Br J Educ Psychol* 2015;85(2):154–71. <https://doi.org/10.1111/bjep.12060>.
- [38] Shanley L. Evaluating longitudinal mathematics achievement growth: modeling and measurement considerations for assessing academic progress. *Educ Res* 2016;45(6):347–57. <https://doi.org/10.3102/0013189X16662461>.
- [39] Pinheiro JC, Bates DM. Linear mixed-effects models: basic concepts and examples. In: *Mixed-effects models in S and S-plus*. Springer-Verlag; 2000. p. 3–56. [https://doi.org/10.1007/0-387-22747-4\\_1](https://doi.org/10.1007/0-387-22747-4_1).
- [40] Hadfield JD. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Software* 2010;33(2). <https://doi.org/10.18637/jss.v033.i02>.
- [41] Hedeker D. Multilevel models for ordinal and nominal variables. In: Leeuw JD, Meijer E, editors. *Handbook of multilevel analysis*. New York: Springer; 2008. p. 237–74. [https://doi.org/10.1007/978-0-387-73186-5\\_6](https://doi.org/10.1007/978-0-387-73186-5_6).
- [42] Ministry of Education. *Scuola, i dati degli scrutini: Alla Secondaria di II grado il 21% degli studenti deve recuperare almeno un'insufficienza, ma diminuiscono le bocciature. Stabili le promozioni nel I grado* 2019. <https://www.miur.gov.it/-/scuola-i-dati-degli-scrutini-alla-secondaria-di-ii-grado-il-21-degli-studenti-deve-recuperare-almeno-un-insufficienza-ma-diminuiscono-le-bocciature-st>.
- [43] INVALSI. Rilevazioni nazionali—I risultati di grado 8: Certificazione delle competenze. [https://INVALSI-areaprove.cineca.it/index.php?get=static&pag=8g\\_descrittori\\_qualitativi](https://INVALSI-areaprove.cineca.it/index.php?get=static&pag=8g_descrittori_qualitativi); 2022.
- [44] Alderighi L, Ballatore RM, Tonello M. Hidden drop-outs: secondary education (unseen) failure in pandemic times. *SSRN Electron J* 2024. <https://doi.org/10.2139/ssrn.4849199>.
- [45] Desimoni M. I livelli per la descrizione degli esiti delle prove invalsi. [https://invalsi-areaprove.cineca.it/docs/2018/Livelli\\_INVALSI\\_g8.pdf](https://invalsi-areaprove.cineca.it/docs/2018/Livelli_INVALSI_g8.pdf); 2018.
- [46] Balfanz R, Herzog L, Mac Iver DJ. Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: early identification and effective interventions. *Educ Psychol* 2007;42(4):223–35. <https://doi.org/10.1080/00461520701621079>.
- [47] Lee VE, Burkam DT. Dropping out of high school: the role of school organization and structure. *Am Educ Res J* 2003;40(2):353–93. <https://doi.org/10.3102/00028312040002353>.
- [48] Hanushek EA. Conceptual and empirical issues in the estimation of educational production functions. *J Hum Resour* 1979;14(3):351. <https://doi.org/10.2307/145575>.
- [49] Contini D, Grand E. On estimating achievement dynamic models from repeated cross sections. *Socio Methods Res* 2017;46(4):988–1017. <https://doi.org/10.1177/0049124115613773>.
- [50] Triventi M, Vlach E, Pini E. Understanding why immigrant children underperform: evidence from Italian compulsory education. *J Ethnic Migrat Stud* 2022;48(10):2324–46. <https://doi.org/10.1080/1369183X.2021.1935656>.
- [51] Alderman H, Behrman JR, Lavy V, Menon R. Child health and school enrollment: a longitudinal analysis. *J Hum Resour* 2001;36(1):185. <https://doi.org/10.2307/3069675>.
- [52] Hirschman C. The educational enrollment of immigrant youth: a test of the segmented-assimilation hypothesis. *Demography* 2001;38(3):317–36. <https://doi.org/10.1353/dem.2001.0028>.
- [53] Agasisti T, Vittadini G. Regional economic disparities as determinants of students' achievement in Italy. *Res Appl Econ* 2012;4(2):p33–54. <https://doi.org/10.5296/rae.v4i2.1316>.
- [54] Argentin G, Triventi M. The North-South divide in school grading standards: new evidence from national assessments of the Italian student population. *Italian J Soc Education* 2015;7(06/2015):157–85. <https://doi.org/10.14658/pupj-ijse-2015-2-7>.
- [55] Bozzano M, Cappelli G, Vasta M. Whither education? The long shadow of pre-unification school systems into Italy's liberal age (1861–1911). *J Econ Hist* 2024;84(1):149–90. <https://doi.org/10.1017/S0022050723000219>.
- [56] Contini D, Tommaso MLD, Mendolia S. The gender gap in mathematics achievement: evidence from Italian data. *Econ Educ Rev* 2017;58:32–42. <https://doi.org/10.1016/j.econedurev.2017.03.001>.
- [57] Agasisti T, Soncin M, Valenti R. School factors helping disadvantaged students to succeed: empirical evidence from four Italian cities. *Policy Stud* 2016;37(2):147–77. <https://doi.org/10.1080/01442872.2015.1127341>.

**Melisa Diaz Lema**, Melisa Diaz is a Junior Assistant Professor at Politecnico di Milano, Department of Management Engineering. Her research falls within the intersection of public management, economics of education, and digital innovation. Her main research interest focuses on the usage of data analytics for decision making in the public sector.

**Chiara Masci**, Chiara Masci is a Junior Assistant Professor at Politecnico di Milano, Department of Mathematics, in the statistical branch. Her research interests are in the development of innovative statistical methods in the area of mixed-effects regression and classification models, both parametric and nonparametric, and in their application in the educational field.

**Mara Soncin**, Mara Soncin is a Junior Assistant Professor at Politecnico di Milano, School of Management. Her research interests are in the educational field, with a focus on (i) digital learning, (ii) school management and (iii) the use of quantitative models for the evaluation of public policies.

**Tommaso Agasisti**, Tommaso Agasisti is a Professor at the School of Management, Politecnico di Milano. He also serves as the Associate Dean for International Relations and Quality at PoliMi GSOM - Graduate School of Management. His research interests lie in the field of Public Economics and Finance, Public Management and Policy, and Public Administration, with a particular focus on schools, universities, and local governments.