



HAL
open science

Teacher quality, family background and student performance in Ecuador

Carlos González-Regalado

► **To cite this version:**

Carlos González-Regalado. Teacher quality, family background and student performance in Ecuador. 2025. <hal-05130196>

HAL Id: hal-05130196

<https://hal.science/hal-05130196v1>

Preprint submitted on 26 Jun 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Teacher quality, family background and student performance in Ecuador

Carlos González-Regalado

Department of Economics, Universidad UTE, Quito, Ecuador

carlosa.gonzalez@ute.edu.ec

June 26, 2025

Abstract

This study investigates the influence of teacher quality and parental education on student performance in standardized tests across four subjects: math, natural science, social science, and language. Using student data from Ecuadorian students between 2015 and 2016, the analysis explores the differential impacts of socioeconomic factors and teaching resources. Findings highlight significant variances in the effects of teacher quality and parental education across subjects, with teacher quality positively influencing outcomes regardless of socioeconomic disparities. Additionally, cognitive skills and access to educational resources show varying levels of importance. The results emphasize the need for policies addressing teacher training and resource allocation.

Keywords: Teacher Quality, Student Achievement, Parental Education, Education Policy.

⁰The author declares no conflicts of interest. Additionally, ChatGPT was used for language and style editing, but all content, analysis, and conclusions are solely the author's responsibility. .

1 Introduction

Education is a cornerstone of economic development that fosters individual opportunity, social mobility, and national growth. Investments in human capital fuel technological progress, improve labor productivity and sustain long-term economic expansion (Krueger, 1999; Mincer, 1974). However, while the broad economic benefits of education are well recognized, identifying the specific mechanisms that drive student achievement remains challenging. A significant obstacle is the reliability of current metrics used to assess educational quality and performance, which often fail to isolate causal effects and introduce substantial measurement errors into empirical research.

A central issue is the inadequacy of existing measures of teacher quality. Many studies rely on proxies such as years of experience, formal qualifications, and class size, yet these indicators capture only partial aspects of teacher effectiveness. Research suggests that these traditional metrics often fail to explain variation in student performance, as they overlook differences in pedagogical skill, subject mastery, and classroom engagement (Hanushek and Rivkin, 2006). Moreover, studies that attempt to estimate teacher effects using value-added models (VAMs) face concerns about bias due to non-random student-teacher assignments and the persistence of unobserved student characteristics. Without a more direct measure of teacher competency, estimates of school quality and student outcomes risk being confounded by selection effects.

This study seeks to improve the measurement of teacher quality by using teacher exam scores as an alternative metric. Unlike conventional indicators, teacher test scores directly assess subject-matter knowledge and pedagogical ability, offering a more reliable predictor of instructional effectiveness. By employing a pseudo-panel data approach, we also mitigate unobserved school-level heterogeneity and enhance causal identification.

Beyond measurement concerns, this study also examines non-linearities in the effects of parental education on student performance. The conventional approach assumes that each additional year of parental schooling has an additive impact on student achievement. However, this assumption may overlook threshold effects: the difference between a parent with a primary education and one with a high school diploma may have a far more significant impact on a child's academic success than incremental increases in years of schooling beyond secondary education. Understanding these nonlinear patterns is critical for shaping policies that effectively reduce educational

disparities.

This study utilizes Ecuadorian educational data for 2015 and 2016, covering both primary and secondary schools. The dataset includes teacher exam scores, student test scores, and school characteristics, allowing us to examine educational outcome variations over time. A key empirical challenge is lacking individual-level panel data to track student progress longitudinally. To address this, we construct a pseudo-panel dataset, aggregating student characteristics at the school level (Deaton, 1985). This technique allows us to control for time-invariant school-level heterogeneity while leveraging Ecuador’s student assignment policy based on residential proximity, reducing self-selection bias.

Both teacher quality and parental education significantly shape student outcomes, with potential nonlinearities in the impact of family background. By employing fixed-effects (FE) and correlated random-effects (CRE) estimators, we mitigate omitted variable bias arising from unobserved cognitive skills, peer effects, and neighborhood influences.

The remainder of this paper is structured as follows: Section 2 surveys the literature of empirical measurement of teacher quality and parental education, Section 3 reviews the methods used to uncover our targeted association, Section 4 presents the dataset, while section 5 introduces the econometric framework. Finally, section 6 summarises our main results and concludes the study.

2 Literature review

2.1 Family background and education

It is widely assumed that each additional year of parental education contributes linearly to student achievement, meaning that higher levels of parental education consistently improve academic performance. However, the existing literature does not examine whether different educational thresholds (e.g., primary vs. secondary education) have varying effects on student outcomes. Despite strong evidence linking parental education to academic success, these studies do not explore whether the impact differs across levels of attainment or whether certain educational milestones (e.g., completing high school vs. university) yield benefits.

Björklund and Salvanes (2010) find that parental education explains approximately 20% of the variation in student schooling, highlighting its substantial role in shaping educational attainment. Similarly, Shoukat et al. (2015) confirm that higher parental education correlates with better academic outcomes, reinforcing the idea that educated parents create more supportive learning environments.

Parental education also plays a significant role in academic involvement and expectations. Toren (2013) distinguish between home-based and school-based parental involvement, showing that more educated parents provide more effective academic support at home. Ali et al. (2022) further emphasize that students whose parents assist with homework and discuss school progress perform better, particularly in households where parents have higher levels of education. Froiland et al. (2013) demonstrate that parental expectations regarding higher education, established early in childhood, have long-term positive effects on student achievement, reinforcing the role of parental education in shaping performance and aspirations.

Zou et al. (2013) highlight cultural variations, showing that Chinese parents with higher education levels are more engaged in their children's academic activities, including supervising schoolwork and investing in additional learning opportunities. Similarly, Stevenson and Baker (1987) find that maternal education is a strong predictor of school involvement, particularly in early childhood. These studies reinforce the positive relationship between parental education and academic success, though they do not examine whether this relationship changes across different educational levels of the parents.

Beyond involvement, parental education influences school-related decision-making and access to academic resources. Educated parents tend to set higher expectations, provide structured learning environments, and seek better educational opportunities for their children. Dookie (2014) and Parveen and Reba (2016) support this argument, finding that students with more educated parents receive greater academic encouragement and support.

While parental education is a strong predictor of student success, it is not the only factor shaping academic performance. Beyond direct educational support, family background influences learning through cognitive development, health, and personality traits. Björklund and Salvanes (2010) highlight that shared household and neighborhood factors play a significant role in determining educational outcomes, indicating that student achievement is shaped by more than just parental schooling. Krapohl et al. (2014)

further emphasize the role of genetic inheritance, showing that intelligence, personality, and psychopathology contribute to student performance. Their findings suggest that cognitive and behavioral traits passed from parents to children may explain part of the link between family background and education outcomes, independent of formal parental involvement.

In addition to genetics, socioeconomic status and parental resources are critical determinants of student success. Yunus and Dahlan (2013) show that higher-income families invest more in their children’s development, providing educational resources that enhance learning opportunities. Cunha and Heckman (2007) support this argument, demonstrating that early childhood investments in cognitive skills yield higher returns than interventions at later stages. These findings indicate that family background shapes academic achievement through multiple channels, including access to educational resources, parental expectations, and cognitive development. While parental education is a crucial factor, it interacts with broader social, economic, and biological influences that contribute to long-term educational outcomes.

2.2 School and teacher quality

The role of school quality in student performance has been widely studied, with early research suggesting that family background plays a more dominant role than school inputs in determining academic achievement (Coleman, 1968). However, later studies recognized the need to account for student composition and selection effects when evaluating school quality (Hægeland et al., 2004). Given that randomized school assignments are rare, most studies rely on observational data and control for individual heterogeneity using cross-sectional or panel data approaches (Krueger, 1999). This methodological challenge complicates efforts to distinguish causal effects from correlations when analyzing school resources and their impact on student learning.

Among school-related factors, teacher quality is widely regarded as one of the most significant determinants of student success, yet it remains difficult to measure. Traditional proxies such as teacher wages, experience, class size, subject specialization, and certifications provide some insights but come with important caveats (Burtless, 2011). Salaries, for instance, do not always reflect instructional ability, as they are often determined by institutional pay structures rather than individual effectiveness. Experience is frequently used as a quality indicator, yet research shows that gains in effectiveness are largest in the first few years of teaching, after which they plateau (Hanushek et al.,

2005). Similarly, class size reductions may improve outcomes for younger students and disadvantaged groups (Nye et al., 2000), but studies have found no consistent long-term impact on achievement as in Hoxby (2000). Given these limitations, there has been increasing interest in performance-based measures of teacher effectiveness, such as value-added models (VAMs), which assess teachers based on their contribution to student learning gains. While these approaches offer a more direct measure of teacher impact, they also face challenges related to measurement error, variability in student test scores, and omitted variable bias.

Beyond observable factors, a substantial portion of teacher quality remains unmeasured, making it difficult to assess what truly defines an effective teacher. Classroom dynamics, instructional methods, and student-teacher interactions are difficult to quantify yet have a significant impact on student learning. Some empirical approaches attempt to estimate unobserved teacher effectiveness using VAMs, which measure a teacher’s contribution to student progress based on improvements in test scores over time.

3 Methodology

This study builds on the identification strategies proposed by Hanushek et al. (2005), Brunello and Checchi (2005), Björklund and Salvanes (2010). By leveraging a panel dataset with two time periods, the analysis exploits temporal variation in educational determinants to estimate their impact on student attainment. The primary estimation equation follows the model proposed by Björklund and Salvanes (2010), which models educational attainment as a function of observed and unobserved heritable endowments. Observed socioeconomic variables that include the availability of technological goods, basic services, household appliances and internet access. Unobserved endowments account for genetic factors and child-rearing practices, captured within the model’s error term.

The model also incorporates variables to control for schooling effects on student attainment, including school quality and its resources, resulting in the following specification:

$$S_{it} = \delta K_{itp} + \gamma_1 X_{ip} + \gamma_2 Z_{itp} + \lambda Q_{its} + \rho R_{its} + \mu_{cit}. \quad (1)$$

Here, K_{itp} represents parental education, which may vary over time, while

Z_{itp} captures time-variant characteristics such as changes in income or resource availability. The term X_{ip} accounts for time-invariant socioeconomic factors, including housing, parental employment, and long-term wealth. χ_{ci} captures other time-invariant factors such as inherited skills and child-rearing behaviors, which could include motivation and goal-setting. Cunha and Heckman (2007) highlight that early investments in childhood skills yield significant impacts on IQ development. Teacher quality, represented by Q_{its} , is measured using standardized exam scores as a proxy for subject knowledge, under the assumption that teachers with subject mastery are more effective educators. Since 2009, the Ecuadorian government has progressively invested in teacher training to enhance both teacher quality and school resources (R_{its}).

Despite these controls, several challenges arise. First, students and teachers could self-select into schools, introducing potential bias. Second, omitted variables, such as evolving relationships between parents and educators, can influence outcomes. Third, behavioral adjustments to policy reforms, such as improved resources, may impact results. For instance, parents might hire private tutors in response to perceived school inadequacies, influencing student outcomes independently of teacher quality.

A panel data framework seems useful to account for unobserved time-invariant characteristics of the students household. Most of the challenges addressed before are significantly associated to long-term wealth or family traditions, that do not vary considerably over time.

3.1 Panel Data Framework

Panel data profits from within-individual variation to identify our parameters of interest, allowing for the control of unobserved heterogeneity, such as family history, without requiring stringent assumptions (Verbeek, 2017). Suppose there are N students indexed by i ($i = 1, \dots, N$) observed across T time periods ($t = 1, \dots, T$). The standard linear model in eq. (1) can be reformulated as:

$$y_{it} = \beta_0 + x_{it}\beta_1 + \epsilon_{it}, \quad (2)$$

where β_0 represents the intercept, x_{it} is the vector of time-varying variables including K_{itp} , Z_{itp} , Q_{sit} , and R_{its} , and β_1 is a set of parameters quanti-

fying their impact. The error term ϵ_{it} comprises an individual-specific error μ_{it} and fixed characteristics (D_{ic}, X_{ip}, J_{is}) associated with the student, household, and school.

Ordinary least squares (OLS) can yield unbiased and efficient estimates if μ_{it} is uncorrelated with x_{it} and the sample is independently and identically distributed. However, unobserved individual effects are often correlated with the observable variables. Additionally, observing the same individuals over time can introduce serial correlation in the error term ϵ_{it} , rendering OLS standard errors unreliable.

To address these challenges, the error term can be expressed as:

$$\epsilon_{it} = \alpha_i + \mu_{it}, \tag{3}$$

where α_i captures all individual-specific, time-invariant factors. While simplifying the model, this assumption is often violated, as Verbeek (2008) highlights. A solution involves including a unique intercept for each individual in the sample, as shown in:

$$y_{it} = \alpha_i + x_{it}\beta + \mu_{it}. \tag{3.7}$$

Here, α_i accounts for both observed and unobserved fixed differences across individuals, with no explicit assumptions about the relationship between x_{it} and α_i . The fixed-effects (FE) method provides two key advantages over cross-sectional models: it increases the efficiency of parameter estimation and mitigates identification issues by accounting for individual-level unobserved heterogeneity (Verbeek, 2017).

However, applying the fixed-effects estimator to eq. (1) eliminates the possibility of identifying time-invariant terms, as they are removed through time-demeaning. While this method addresses omitted variable bias related to time-invariant characteristics, such as individual skills or permanent wealth, it cannot estimate their effects on educational attainment. Additionally, the coefficients for time-variant variables, such as parental education and wealth, are identifiable only if there is sufficient within-individual variation after the time-demeaning process. In such case, alternative methodologies may be necessary to fully capture the parameters of interest in this study.

Our dataset structure requires a model capable of efficiently addressing OLS shortcomings, but allowing to estimate the effect of time-invariant

variables. Among the available options, the correlated random effects model (CRE) is deemed most suitable for estimating the parameters in eq. (1).

Chamberlain (1982) proposed an alternative that models the conditional expectation of unobserved time-invariant heterogeneity as a linear function of the observed time-invariant covariates:

$$\alpha_c = \bar{x}_c \gamma + \eta_c. \quad (4)$$

Equation (4) expresses a linear relationship between the unobserved individual effects α_c and the time-averages of the covariates \bar{x}_c . Here, η_c represents a true random effect, maintaining the assumptions of a standard random effects model. Our setup is supported by Cunha and Heckman (2007), who indicate that parental investments during early childhood substantially enhance cognitive skills compared to children from less affluent backgrounds. Consequently, unobserved cognitive abilities, such as intelligence, are associated with permanent family wealth levels.

It is essential to account for the assumption of strict exogeneity, i.e., the *true* random effect term is not correlated with the observed variables. This assumption is unlikely to be violated because of our cohort average estimation that is introduced in the next section. The fluctuations in the average cohort exam grades do not influence subsequent period performances, as cohort averages are calculated using new individuals. Consequently, the CRE model ensures that the coefficients for time-variant variables are consistent with those obtained through fixed effects estimation Wooldridge (2013).

3.1.1 Cluster Averaging

Specifically, the INEVAL dataset records students in each grade only once per examination period, preventing the tracking of individual students over time while maintaining consistent exam complexity. Such tracking is crucial for certain estimators (e.g., the within-estimator) to account for unobserved individual time-invariant heterogeneity.

To address this issue, this study employs the pseudo panel data approach introduced by Deaton (1985). This approach estimates cohort means for individuals and utilizes these averages over time to construct a panel dataset. Repeated cross-sections from INEVAL data between 2015 and 2016 facilitate the creation of a school-level dataset. However, accurate estimation of cohort

presents a significant challenge; inaccuracies can bias the estimates of our estimators. Additionally, the averaging process significantly reduces within-cohort variation.

Deaton (1985) suggests that cohort averages can provide consistent estimates of β in eq. (3.7) when repeated cross-sections are available, even if unobserved heterogeneity is correlated with explanatory variables. This requires defining C non-overlapping cohorts where each observation belongs to a fixed cohort across all time periods. For example, all 13th-grade students in 2015 would form a single cohort. In this study, time-invariant criteria, such as grade or school, are used to define cohorts, ensuring all cohort mean covariates are observable for all individuals.

After clustering individual observations into cohort averages, eq. (3.7) can be rewritten as:

$$y_{ct} = x_{ct}\beta + \alpha_{ct} + \mu_{ct}, \quad (5)$$

where $c = 1, \dots, C$ and $t = 1, \dots, T$. Here, \bar{y}_{ct} represents the average outcome variable for cohort c in time t , and similar notation applies to other model variables. The resulting model functions as a pseudo panel with T time periods and C observations. However, α_{ct} is likely time-dependent and correlated with \bar{x}_{ct} , leading to biased parameter estimates if OLS is used with the composite error term $\alpha_{ct} + \mu_{ct}$. With sufficiently large cohorts, α_{ct} converges to the true cohort effect (Verbeek, 2017). Assuming α_{ct} is fixed over time ($\alpha_{ct} = \bar{\alpha}_c$), the within-estimator for eq. (5) can be expressed as:

$$\hat{\beta} = \left(\sum_{c=1}^C (x_{ct} - \bar{x}_c)(x_{ct} - \bar{x}_c)' \right)^{-1} \sum_{c=1}^C (x_{ct} - \bar{x}_c)(y_{ct} - \bar{y}_c), \quad (6)$$

where \bar{x}_c and \bar{y}_c represent the time averages for each cohort c . The properties of these estimators depend on the asymptotics employed. One approach assumes $N \rightarrow \infty$ with C fixed, ensuring $n_c \rightarrow \infty$. This results in genuine time variation for cohort averages, provided cohorts are appropriately constructed. Measurement errors in cohort averages diminish with larger cohort sizes. It is assumed that these errors have zero mean and are independent of the true variable values.

The INEVAL dataset provides a sufficient number of individuals per time period for consistent cohort average estimation. However, there is a trade-

off between the accuracy of cohort mean estimates and the within-cohort variation. Despite school clustering offering sufficient within-variation for identifying parental education and school quality effects, these variables exhibit limited time variation, therefore, our estimates should be estimated with caution.

4 Data

4.1 General Description of the Survey Data

The dataset for this study was developed and implemented by INEVAL, adhering to the standards outlined by Rutkowski and Rutkowski (2013a). It includes information on standardized examination scores for students at the end of each educational level (4th, 7th, 10th, and 13th grades). These exams cover mathematics, natural sciences, social studies, and languages, progressively increasing in complexity to evaluate educational quality and align national achievement with international benchmarks. For the 13th grade in 2015, 237,186 students from 3,268 schools participated, while in 2016, 247,682 students from 3,179 schools were evaluated. The sample size for the 13th-grade examination approximates 35%. However, due to incomplete socioeconomic questionnaires, not all data was available for this study. This attrition effect is analyzed later in this section.

In 2016, INEVAL also administered subject-specific tests to 98,921 teachers from 13,842 schools nationwide. These tests included socioeconomic and school-related data, with infrastructure, classroom, and laboratory quality derived from teacher surveys. Notably, this study excludes teacher-level socioeconomic variables from the analysis of student attainment. Teacher test scores ranged from 0 to 1000, similar to the student assessments, but due to the lack of direct teacher-student matching records, individual observations were omitted. Instead, teacher grades were averaged at the school level without additional weighting.

The data was collected through random cross-sectional sampling for each examination period. Two periods (2015 and 2016) are available for 13th-grade students, while an additional year (2014) is available for other grades.

4.2 Summary Statistics

Table 1 summarizes the number of students, average scores, and key covariates at the school level for 13th-grade students. Student data was aggregated at the school level for the identification strategy. Despite some missing socioeconomic survey data, school averages remained unaffected. The dataset includes 7,316 schools across two years, with the balanced panel comprising 6,962 observations from 3,481 schools between 2015 and 2016. No significant differences were observed in the main outcome variables between the full and balanced samples, with data loss remaining under 5%. Consequently, the information loss across schools and time periods is assumed to be random. Notably, only 2% of schools in the sample are private, while the majority belong to the public education system.

Cohort averaging requires a substantial number of observations within each category. The mean cohort size is 88.27 students, with a standard deviation of 108.45, suggesting potential variability in cohort sizes. However, tests of sample means for key covariates across provinces and years revealed no significant differences, indicating that cohort size does not pose a problem for this analysis. Additionally, the data showed no systematic differences in educational outcomes between males and females. Therefore, cohorts at the school level were calculated separately for males and females.

The final student test score is derived from the unweighted average of the test's four subsections. Each subsection is graded on a scale from 0 to 1000 points, with a minimum achievable score of 400 points and increments of one point. Nationwide, the average grade for the entire sample was 746.4 points. Among the four examinations, mathematics had the lowest average score of 698.8 points, with a standard deviation of 97.34.

Social studies scores had the highest average at 805.9 points. Students in 4th grade achieved better results across all subjects compared to 7th and 10th grades. However, scores generally improved by the 13th grade, except in language, which remained relatively unchanged. This improvement for senior year students could be attributed to specific preparation programs for senior students. Briones Rugel et al. (2011) noted that students performed better in subjects other than mathematics in 2009 compared to 2016.

Ecuador is geographically and culturally diverse, with four distinct regions that influence educational outcomes, as observed by Briones Rugel et

Table 1: Summary Statistics of Main Variables Balanced Panel

	Obs	Mean	SD	Min.Val	Max.Val
Examinations					
Average (avg)	6962	746.4	81.45	515.1	996.9
Math	6962	698.8	97.34	443	998.7
Language	6962	755.2	76.99	500	996.2
Natural	6962	725.8	99.11	443	1000
Social	6962	805.9	93.12	500	998.3
Socioeconomic Indicators					
Fam Tech	6962	3.660	0.655	0	5
Fam Goods	6962	1.167	0.709	0	4
Fam Services	6962	3.231	0.807	0	5
Fam Books	6962	3.499	0.415	1	5
Parent Education	6962	9.645	1.961	0	17
Mother Education	6962	9.828	2.251	0	17
Father Education	6962	9.462	1.993	0	17
Socioecon_index	6962	-0.423	0.624	-2.365	2.080
School Quality Indicators					
Teacher Quality	6962	663.3	38.61	324	890
Tech_school	6962	3.388	0.925	0	4
Infra_score	6962	3.920	1.545	0	7
Sch_room	6962	3.156	0.461	1	4

al. (2011). Table 2 highlights these regional differences. The Galapagos and Andean regions report higher average performance and socioeconomic indicators, while parents in the Coast and Eastern regions tend to have more years of education on average. Teacher quality is comparable across the Coast, Eastern, and Galapagos regions but significantly higher in the Andes. These regional variations necessitate the inclusion of regional controls in the model specification. Additional district-level controls further account for localized geographic and administrative effects on educational outcomes.

In 2015, the national examination average was 751.3 points, dropping to 746.4 in 2016, likely due to the major earthquake in Ecuador (april, 2016). Table 3 compares the averages of key variables between the two years. A minor decrease of about 0.2 years in parental education is observed, which may result from sampling differences. Teacher scores remained stable, indicating

Table 2: Summary Statistics of Main Variables by Region

Region	G.Average	P.Year.Educ	Grade.Teach	Econ.Index
Coast	729.0	9.912	648.8	-0.482
Galapagos	763.0	11.05	649.5	0.317
Eastern	698.5	9.623	646.0	-0.619
Andes	772.1	9.410	680.2	-0.332
Total	746.4	9.645	663.3	-0.423

that the sampling process did not compromise school quality balance. The socioeconomic index declined slightly, reflecting Ecuador’s economic contraction in 2016 due to falling oil prices and government budget challenges.

Table 3: Mean of variables across time periods

	G.Average	P.Year.Educ	Grade.Tech	Econ.Index
Year 2015	751.3	9.746	663.3	-0.430
Year 2016	741.5	9.545	663.4	-0.417
Total	746.4	9.645	663.3	-0.423

4.3 Main variables

The construction of the main covariates in this study adheres to international standards outlined by INEVAL (2017), Rutkowski and Rutkowski (2013b), and Flotts et al. (2015). The first set of variables pertains to family background characteristics, grouped into four categories: parental education levels, technological resources available in the household, number of books at home, and the availability of basic household goods. Due to data limitations, parental occupation was not included in the analysis. Another set of variables measures household income, services, and goods.

Parental Education is defined as the average number of completed years of education of the father and mother. The education system categorizes the first seven years as primary education, the next six years as secondary education, and most university programs span four years, making the variable range from 0 to 17 years.

Technological Goods Availability is constructed based on the availability of a DVD player, computer, cell phone, TV, and desk at the household

level, resulting in a score ranging from 0 to 5. *Family Goods* include items essential for academic success, such as a bed, car, washing machine, and microwave. *Household Services*, which influence student outcomes, include access to running water, sewerage, electricity, a fixed phone line, and cable TV. The *Books Availability* variable encompasses the presence of a dictionary and other educational books. These variables, like technological availability, are constructed by summing the individual components. Additionally, Internet service is analyzed separately from household services.

Since household income data is unavailable in the INEVAL dataset, this study uses a *Socioeconomic Index* developed by INEVAL (2017). This index, derived through factor analysis and principal component analysis (PCA), captures the variance across various household wealth-related covariates. Despite being correlated with family background variables, it serves as a reliable proxy for the missing income data. This index is crucial in the model specification to account for time-invariant wealth.

Teacher quality, a central variable in this study, is measured through teacher exam scores related to the specific content of their assigned subjects. The *School Quality Variables*, are part of a survey conducted by the Ministry of Education in 2013.

The *School Room Variable* measures the availability of items such as electronic blackboards, teacher desks, and teaching guides, ranging from 0 to 4. The *Technological Facilities Variable* evaluates the presence of a computer lab, projector, audiovisual lab, and chemistry lab, scoring up to 4. Finally, the *Infrastructure Variable* assesses facilities such as sports grounds, dining areas, administrative offices, teacher rooms, art and music rooms, and teacher-use computers, with a maximum score of 7.

5 The Education system in Ecuador

The Ecuadorian education system, as established until 2016, required universal and compulsory education for all students in the 1st to 13th grade. The Ecuadorian Constitution guarantees access to free public education, encompassing tuition and ancillary services such as breakfast, uniforms, and textbooks. The government sought to ensure that quality education was strategically placed and accessible throughout the nation.

General Basic Education (GBE) was segmented into four academic phases.

The Preparatory level encompassed 1st grade, intended for children up to five years of age. The Basic Elemental Education phase encompassed grades 2 through 4, succeeded by Basic Middle Education for grades 5 to 7. The last phase, Basic Upper Education, encompassed grades 8 through 10. At the conclusion of each stage, students were mandated to undertake standardized assessments in mathematics, language, natural sciences, and social studies to evaluate their academic advancement.

The national curriculum encompassed language and literature, English, mathematics, natural sciences, social sciences, physical education, culture, and the arts. In 2015, a significant education reform reconfigured the system, supplanting the prior model that permitted students in their final years to concentrate in subjects corresponding to their professional ambitions. In the previous system, students studying engineering or mathematics engaged in an advanced curriculum of mathematics and physics, whilst those focused on medical or life sciences undertook courses in chemistry and biology. Subsequent to the reform, all students were provided with a same general education, irrespective of their prospective university major.

Student placement in public schools was based on geographic proximity and meritocratic standards. Enrollment occurred biannually, in March and September. For first-grade enrollment, school allocations favored proximity; if a school reached full capacity, kids were assigned to the next available institution. Academic performance was also a factor at the eighth-grade level. Students with the highest elementary school grades were prioritized in selecting their chosen schools, while the remaining vacancies were allocated based on proximity and academic performance averages. Individuals who failed to obtain a position at their desired university were sent to alternative colleges according to availability.

To standardize educational quality, MINEDUC created a framework that delineates requirements for school management, teacher efficacy, student performance, and infrastructure. Administrative standards emphasized resource management, financial planning, and institutional governance to ensure efficient school operations and financial stability. Schools were urged to establish dispute resolution methods, participate in community activities, and prepare contingency plans for crises. Faculty and staff were mandated to comply with codes of behavior, and schools were obligated to offer ongoing training programs for educators.

Teacher performance requirements prioritized topic expertise, lesson design, and classroom administration. Educators were required to exhibit ex-

expertise in their subject areas and comprehend the interconnections among other disciplines. Lesson planning must be systematic, including suitable instructional materials and resources. Furthermore, educators were mandated to participate in ongoing professional development and sustain a learning atmosphere that promoted student involvement and discipline.

Student performance was evaluated according to learning outcomes in fundamental areas. In language and literature, pupils advanced from acquiring fundamental oral communication abilities in early grades to crafting intricate written assessments of scientific and literary works in their last years of high school. The mathematics curriculum encompassed subjects from numerical operations and fundamental geometry in early grades to statistics, probability, and advanced algebra in later grades. In social studies, pupils initially studied family and civic obligations, subsequently examining Ecuador's historical history. The natural sciences curriculum familiarized pupils with ecosystem dynamics, biological systems, and the principles of physics and chemistry, advancing from basic concepts in early school to scientific modeling and hypothesis testing at higher levels.

Infrastructure standards guaranteed that educational institutions offered an ideal learning environment. Schools were mandated to provide sufficient classrooms, scientific laboratories, libraries, and athletic facilities. In several instances, libraries, internet laboratories, and sporting facilities were constructed for communal utilization. These infrastructural principles aimed to ensure equitable access to educational resources for all students, irrespective of their geographical location.

By 2016, the Ecuadorian education system functioned under established norms to enhance quality and accessibility. The incorporation of teacher competency evaluations, systematic student assessments, and enhancements to infrastructure demonstrated an initiative to establish a more equal and efficient national education system.

6 Empirical Results

6.1 Effects of School Quality and Family Background on Educational Attainment

Our identification strategy enables the analysis of the relationship between parental education, school quality, and student performance in the 13th grade. A fundamental assumption of the strategy in eq. (7) is the absence of correlation between unobserved time-invariant variables, the observed regressors, and the error term. For this assumption to hold, the observation of household income would be necessary, but it is not available in the INEVAL dataset.

$$y_{it} = \delta K_{itp} + \lambda Z_{itp} + \gamma X_{ip} + \rho R_{it} + \chi_{ci} + \phi_i + \tau_t + \mu_{it}, \quad (7)$$

In this equation:

- K_{itp} : Parental education level.
- Z_{itp} : Time-variant characteristics influencing educational outcomes.
- X_{ip} : Time-invariant socioeconomic wealth and household characteristics.
- χ_{ci} : Other unobserved time-invariant factors, such as inherited skills and child-rearing practices.
- Q_{sit} : Teacher quality, measured through standardized teacher examinations.
- R_{its} : School time-variant quality and resources.
- ϕ_i : Time-invariant school attributes, such as infrastructure.

To simplify the interpretation and presentation of the results, the variables in Equation (7) are reorganized as shown in eq. (8):

$$y_{it} = \delta K_{itp} + \gamma X_{ip} + \lambda Z_{it} + \rho R_i + \tau_t + \mu_{it}, \quad (8)$$

Here, the variables are grouped as follows:

- **Family Background:** Includes observed time-variant parental education (δK_{itp}) and time-invariant wealth (γX_i).
- **Composite Error Term (μ_{it}):** Captures unobserved factors like cognitive skills, motivation, peer and sibling effects, and social networks.
- **Teacher Quality (λZ_{it}):** Based on teacher examination scores.
- **Time-Invariant School Features (ρR_i):** Encompasses infrastructure and resources, which are treated as constant due to limited time observations (T).
- **Common Time Effects (τ_t):** Reflects effects shared across individuals, such as special programs or enhanced school management.

The assumption of small T limits the variability across time, leading to the treatment of school resources and infrastructure as time-invariant. Additionally, special courses or improved school management are accounted in the composite error term (μ_{it}).

6.2 Cohort Averages and Panel Data Framework Results

Table 4 introduces four estimation procedures to verify whether our estimates are driven by the particular characteristics of them or the estimates hold across specifications, i.e., pooled OLS estimation, followed by FE, RE, and CRE models.

Due to the fact that many observed features are correlated with the composite error term, it is necessary to compare the OLS estimates with the panel data ones. The bias could either be positive or negative according to the context. On the one hand, self-selection of students into high-quality schools or those with high-quality teachers could create an upward bias. Conversely, government resource redistribution to less advantaged students might cause a downward bias, as these students receive more support. A test for the joint significance of geographic and time effects rejects the null hypothesis that their coefficients are zero.

In table 4, teacher quality, parental education, and geographic controls (province and canton) explain 47% of the variance in student grades on IN-EVAL exams, leaving significant heterogeneity unexplained. Specification

Table 4: Regression Estimates Model Comparison

	(1) OLS	(2) FE	(3) RE	(4) CRE
ParentEducation	4.524*** (0.968)	1.799* (1.039)	4.293*** (0.930)	3.195*** (0.888)
Internet	61.88*** (7.234)	30.03*** (9.549)	60.46*** (7.204)	46.00*** (7.377)
FamTech	-4.079 (2.728)	-0.494 (3.290)	-3.179 (2.597)	-6.268** (2.673)
FamBooks	25.84*** (3.156)	10.53*** (3.945)	22.30*** (3.040)	10.02*** (3.167)
FamGoods	14.65*** (3.078)	-5.397 (3.716)	11.96*** (2.954)	7.194** (3.044)
TeacherQuality	0.230*** (0.0487)	0.244*** (0.0492)	0.217*** (0.0477)	0.217*** (0.0477)
FemaleTeacher	20.61*** (7.633)	22.23*** (7.720)	19.64*** (7.699)	19.64*** (7.699)
Infrastructure	2.573** (1.002)	3.019*** (1.015)	2.481** (0.985)	2.481** (0.985)
Morning	14.56** (6.177)	14.69** (6.214)	12.13* (6.198)	12.13* (6.198)
D.2016	-11.31*** (2.084)	-9.974*** (2.006)	-11.04*** (2.070)	-10.18*** (2.040)
D.Coast	21.60*** (7.306)	21.56*** (7.291)	22.78*** (7.456)	22.78*** (7.456)
Constant	300.9*** (39.80)	694.2*** (16.73)	280.4*** (35.78)	216.0*** (36.02)
Observations	6962	6962	6962	6962
R^2	0.471	0.031	0.471	0.480
R^2 Within	-	0.0313	0.0221	0.0260
R^2 Between	0.245	-	0.582	0.587
R^2 Overall	0.170	-	0.471	0.480
Het. Variance	69.32	34.06	33.80	33.80
Hausman Test χ^2	156.1	55.72	-	-
Hausman Test p-val	0.000	0.000	-	-
Breusch-Pagan LM χ^2	251.70	261.45	-	-
Breusch-Pagan LM p-val	0.000	0.000	-	-
R^2 Resid OLS	0.209	-	0.002	0.000
CRE p-val	-	-	-	0.000
Controls				
Canton	Yes	No	Yes	Yes
Province	Yes	No	Yes	Yes
Correl RE	No	No	Yes	Yes

Parroquia Clustered Standard Errors in Parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(1) indicates upward-biased estimates for parental education, consistent with Björklund and Salvanes (2010), where one additional year of parental education increases student scores by 4.5 points. However, this effect decreases to 1.8 points (significant at the 10% level) in the FE specification (2). Specifications (3) and (4) (RE and CRE models) produce estimates similar to OLS but are inconsistent according to the Hausman test. FE estimates remain consistent if time-variant regressors are uncorrelated with the idiosyncratic error term. The Breusch-Pagan test rejects the null hypothesis that the variance of time-invariant heterogeneity equals zero, suggesting that income-related variables inadequately capture unobserved individual effects.

Teacher quality consistently improves student performance across all specifications, with each one-point increase in teacher exam scores raising student scores by 0.21 to 0.24 points (significant at the 1% level). The infrastructure index also shows consistent positive effects, with better facilities (e.g., sports halls, classrooms, and libraries) improving performance by up to 3 points for every one-point increase in the index. Additionally, increasing the proportion of female teachers by 1% correlates with a 20-point increase in student scores, likely due to the higher participation of skilled women in education. Morning class attendance improves performance by 12.1 to 14.6 points compared to evening or night classes.

Students with internet access achieve, on average, more than 60 points higher in specifications (1) and (3), while the effect decreases to 30 points in the FE specification (2), suggesting potential upward bias. Wealthier households are more likely to provide internet access and better-quality education. Surprisingly, technological goods like cell phones and televisions are insignificant in specifications (1) to (3), but specification (4) shows a negative effect at the 5% significance level. As expected, book availability at home positively impacts performance, increasing scores by 25.8 points. This effect decreases to 10.5 points in FE and CRE specifications.

6.3 Socioeconomic Measure Impact

Variables such as access to household goods, services, and technology do not fully capture the effect of wealth or income in the student attainment model specified in eq. (1). Table 5 presents results that incorporate the INEVAL socioeconomic index.

After adding the socioeconomic index, table 5 shows a significant reduc-

Table 5: Regression Estimates Model Comparison + Interactions

	(1) OLS	(2) FE	(3) RE	(4) CRE
ParentEducation	-21.75*** (3.169)	-10.24** (4.080)	-20.41*** (3.085)	-8.694** (3.392)
ParentEducSQ	1.432*** (0.179)	0.661*** (0.234)	1.335*** (0.175)	0.580*** (0.194)
socioecon_index	19.88** (8.124)	8.834 (9.532)	21.42*** (7.695)	5.242 (7.952)
iec*D.2016	12.05*** (3.073)	13.41*** (3.140)	12.69*** (3.097)	12.95*** (3.128)
Internet	30.21*** (8.994)	13.09 (11.52)	27.36*** (8.878)	16.67* (9.404)
FamTech	-4.125 (3.025)	-1.943 (3.733)	-4.751 (2.913)	-0.889 (2.984)
FamGoods	1.186 (3.560)	-7.830* (4.134)	-0.159 (3.377)	-6.146* (3.457)
FamServices	-1.242 (2.651)	-0.582 (3.381)	-0.719 (2.485)	-0.297 (2.826)
FamBooks	21.55*** (3.074)	10.87*** (3.902)	19.18*** (2.968)	10.82*** (3.101)
TeacherQuality	0.220*** (0.0483)	0.230*** (0.0485)	0.203*** (0.0484)	0.203*** (0.0484)
FemaleTeacher	21.35*** (7.355)	22.20*** (7.383)	21.35*** (7.468)	19.34*** (7.468)
Morning	7.495 (6.135)	7.790 (6.153)	7.495 (6.142)	5.395 (6.142)
D.2016	-4.728** (2.095)	-3.774* (2.099)	-4.351** (2.090)	-4.314** (2.106)
D.Coast	22.53*** (7.236)		22.47*** (7.212)	24.02*** (7.696)
Constant	520.7*** (45.63)	764.8*** (32.32)	459.8*** (45.45)	443.9*** (70.87)
Observations	6962	6962	6962	6962
R^2	0.491	0.049	0.490	0.502
R^2 Within		0.0492	0.0410	0.0492
R^2 Between		0.286	0.599	0.607
Het. Variance	67.17	33.05	32.63	32.63
Hausman Test χ^2		108.6		7.297
Hausman Test p-val		0.000		0.697
Breusch Pagan LM χ^2			264.65	274.28
Breusch Pagan LM p-val			0.000	0.000
R^2 Resid OLS	0.118		0.022	
CRE χ^2				97.56
CRE p-val				0.000
Controls				
Canton	Yes	No	Yes	Yes
Province	Yes	No	Yes	Yes
Correl RE	No	No	No	Yes

Parroquia Clustered Standard Errors in Parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

tion in the magnitude and significance of family household coefficients, as expected, particularly the parental education estimate. The impact drops to less than one point for each additional year of parental education. Two possible explanations arise: (1) Wealth levels in Ecuador may better explain children’s average school exam performance, aligning with Björklund and Salvanes, 2010, who argue that sibling and neighborhood effects outweigh parental education in explaining educational attainment or (2) The socioeconomic index is highly correlated with parental education, potentially absorbing its variance in the model or introducing a multicollinearity issue in the specification. Notably, the socioeconomic index is significant at the 5% level across specifications (1) to (4), with its coefficient halving under the FE estimator in specifications (2) and (4). This suggests that financial stability, such as home ownership or savings, creates a conducive environment for learning. The joint significance test for CRE time averages rejects the hypothesis that the coefficients are zero, and the Hausman test indicates no systematic differences between FE and CRE coefficients (p -value = 0.75).

The teacher quality coefficient remains robust across all specifications in Table 5.2, with a slight decrease from 0.218 in OLS (specification 1) to 0.196 in CRE (specification 4). All coefficients are significant at the 1% level, confirming the accuracy of these estimates. The proportion of female teachers also shows a small reduction compared to table 4. For the school infrastructure index, the OLS coefficient (1) remains consistent across Tables 5.1 and 5.2, but the CRE coefficient (4) decreases from 2.8 to 1.64 points per unit increase, both significant at the 10% level. This suggests that the initial coefficients in Table 5.1 could be upwardly biased.

6.4 Non-linear effects of parental education

To investigate further the association between parental background on exam results, eq. (1) account for non-linear effects of parent education. When the square term of parental education is included, both the linear and squared terms are significant at the 5% level. This finding implies that the effect of parental education differs between low- and high-education households. In specification OLS (1) of table 6, lower levels of parental education negatively impact children’s attainment. The net effect is calculated as follows:

$$M_{OLS} = -27.75 + 2 \cdot 1.432 \cdot \text{ParentYearsEduc}$$

Table 6: CRE Estimates for Exam Subjects

	(1) Avg FS	(2) Math	(3) N. Science	(4) Social E.	(5) Language
ParentEducation	-8.694** (3.392)	-3.435 (5.348)	-13.66*** (3.751)	-8.693** (4.608)	-8.015** (4.052)
ParentEducSQ	0.580*** (0.194)	0.236 (0.300)	0.811*** (0.224)	0.482* (0.255)	0.734*** (0.228)
socioecon_index	5.242 (7.952)	-2.740 (10.93)	10.71 (11.26)	19.34* (10.76)	-7.870 (8.655)
iec*D.2016	12.95*** (3.128)	10.29** (4.075)	22.96*** (4.308)	7.755** (3.690)	10.50*** (3.062)
Internet	16.67* (9.404)	20.81** (11.94)	9.197 (12.23)	-3.248 (12.74)	40.97*** (9.296)
FamBooks	10.82*** (3.101)	9.951** (4.106)	9.033** (4.095)	13.07*** (4.063)	11.18*** (3.007)
TeacherQuality	0.203*** (0.0484)	0.168*** (0.0597)	0.219*** (0.0530)	0.234*** (0.0548)	0.190*** (0.0505)
FemaleTeacher	19.34*** (7.468)	15.51* (9.157)	26.57*** (9.029)	19.11*** (9.183)	16.25*** (6.630)
Infrastructure	1.345 (0.979)	1.597 (1.209)	1.404 (1.217)	1.414 (1.187)	0.955 (0.902)
Morning	5.395 (6.142)	7.012 (7.322)	7.249 (7.630)	4.948 (7.079)	2.305 (5.318)
D.2016	-4.314** (2.106)	-3.143 (2.583)	-11.04*** (2.825)	18.03*** (2.225)	-21.38*** (2.360)
D.Coast	24.02*** (7.696)	3.325 (10.92)	19.00** (9.407)	24.20*** (8.584)	49.21*** (8.645)
Constant	443.9*** (70.87)	479.5*** (82.95)	366.7*** (83.69)	376.7*** (86.01)	553.8*** (64.59)
Observations	6962	6962	6962	6962	6962
R^2 Within	0.049	0.015	0.076	0.044	0.173
R^2 Between	0.607	0.559	0.533	0.525	0.608
R^2 Overall	0.502	0.437	0.417	0.409	0.501
Het. Variance	32.63	36.99	36.11	35.92	25.56
Controls					
Canton	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes
Correl RE	Yes	Yes	Yes	Yes	Yes

Parroquia Clustered Standard Errors in Parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For households with an average of 7.59 years of education, the net effect is zero, with higher levels of education increasing student performance and lower levels reducing it. While the OLS estimates suggest bias, the FE specification (2) accounts for the bias but cannot control for time-invariant variables. The CRE model (3) provides our preferred estimates, with the net effect calculated as follows:

$$M_{OLS} = -8.694 + 2 \cdot 0.58 \cdot \text{ParentYearsEduc}$$

Households with more than 7.50 years of education report positive effects on children’s grades. This estimate is slightly lower than in specification (1) but not substantially different. Teacher quality remains significant at the 1% level in CRE (3), with consistent coefficients compared to table 4, indicating robust measurement. Residual analysis using OLS regressions for specifications (2) to (4) reveals no unobserved correlation, with specification (4) showing an R^2 of zero.

As noted by Ferguson et al. (2007), poverty and low education levels often lead to suboptimal parental decisions, affecting long-term outcomes like dietary choices and school readiness. When controlling for the nonlinear term, the socioeconomic index in specifications (2) and (4) becomes insignificant, suggesting that time-variant economic conditions have a limited impact on average student achievement after controlling for parent education variation.

6.5 Differences Across Subjects in the General Exam

The INEVAL standardized exam evaluates children’s knowledge in four subjects: math, natural sciences, social sciences, and language (Spanish). While the tests maintain a consistent level of complexity across years, the students’ performance on each subject depends greatly on their attitudes and skills. The corresponding estimates are presented in table 7.

Table 7, specification (1), indicates that for parental education to have a positive effect, the average years of education of parents should be at least 7.5 years. However, in math (2), neither the square of parental years of education nor the parental education coefficient is significant at the 10% level, and both have lower magnitudes. This suggests that, within the INEVAL evaluated students, unobserved individual skills play a more significant role in determining math performance. Teacher quality remains a significant and

Table 7: CRE Estimates for Exam Subjects

	(1) Avg FS	(2) Math	(3) N. Science	(4) Social E.	(5) Language
ParentEducation	-8.694** (3.392)	-3.435 (5.348)	-13.66*** (3.751)	-8.693** (4.608)	-8.015** (4.052)
ParentEducSQ	0.580*** (0.194)	0.236 (0.300)	0.811*** (0.224)	0.482* (0.255)	0.734*** (0.228)
socioecon_index	5.242 (7.952)	-2.740 (10.93)	10.71 (11.26)	19.34* (10.76)	-7.870 (8.655)
iec*D.2016	12.95*** (3.128)	10.29** (4.075)	22.96*** (4.308)	7.755** (3.690)	10.50*** (3.062)
Internet	16.67* (9.404)	20.81** (11.94)	9.197 (12.23)	-3.248 (12.74)	40.97*** (9.296)
FamBooks	10.82*** (3.101)	9.951** (4.106)	9.033** (4.095)	13.07*** (4.063)	11.18*** (3.007)
TeacherQuality	0.203*** (0.0484)	0.168*** (0.0597)	0.219*** (0.0530)	0.234*** (0.0548)	0.190*** (0.0505)
FemaleTeacher	19.34*** (7.468)	15.51* (9.157)	26.57*** (9.029)	19.11*** (9.183)	16.25*** (6.630)
Infrastructure	1.345 (0.979)	1.597 (1.209)	1.404 (1.217)	1.414 (1.187)	0.955 (0.902)
Morning	5.395 (6.142)	7.012 (7.322)	7.249 (7.630)	4.948 (7.079)	2.305 (5.318)
D.2016	-4.314** (2.106)	-3.143 (2.583)	-11.04*** (2.825)	18.03*** (2.225)	-21.38*** (2.360)
D.Coast	24.02*** (7.696)	3.325 (10.92)	19.00** (9.407)	24.20*** (8.584)	49.21*** (8.645)
Constant	443.9*** (70.87)	479.5*** (82.95)	366.7*** (83.69)	376.7*** (86.01)	553.8*** (64.59)
Observations	6962	6962	6962	6962	6962
R^2 Within	0.049	0.015	0.076	0.044	0.173
R^2 Between	0.607	0.559	0.533	0.525	0.608
R^2 Overall	0.502	0.437	0.417	0.409	0.501
Het. Variance	32.63	36.99	36.11	35.92	25.56
Controls					
Canton	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes
Correl RE	Yes	Yes	Yes	Yes	Yes

Parroquia Clustered Standard Errors in Parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

positive determinant for math scores, though its effect is weaker compared to other subjects.

For natural sciences, parental education has a stronger impact, requiring an average of 8.4 years of parental schooling to yield a positive effect—higher than the threshold for the overall score. Meanwhile, for language, the threshold is the lowest at 5.5 years. Teacher quality is consistently significant at the 1% confidence level across all subjects. Notably, in social studies and language, both the square of parental years of education and the parental education coefficient are significant at the 10% level.

7 Conclusion

This study provides robust evidence of the causal effects of parental education and teacher quality on student performance in standardized exams within the Ecuadorian educational system. By employing a correlated-random effects (CRE) model and leveraging pseudo-panel data constructed from IN-EVAL’s dataset (2015–2016), we effectively addressed challenges related to unobserved heterogeneity. The findings underline the importance of family background and teaching quality in shaping educational outcomes, although these effects vary by student age, socioeconomic context, and subject area.

One key finding is that parental education exerts a nonlinear influence on student performance. While parents with more than 7.5 years of schooling positively contribute to their children’s grades, those with fewer years of education have a detrimental effect. This finding highlights the need for targeted interventions to support low-educated households to mitigate adverse outcomes. Similarly, teacher quality consistently showed a significant positive impact, particularly for senior-year students, where every additional point in teacher examination scores corresponded to a grade increase between 0.199 and 0.230 points. These findings suggest that a joint effort by parents, guardians, and teachers is crucial to improving student performance.

Socioeconomic factors also play a pivotal role, particularly in rural areas and subjects like social studies and languages. Wealth enhances access to educational materials, such as books and the Internet, which consistently improve exam grades. However, the effects of other goods and services in the household, though significant in some models, lack consistency. Additionally, while school infrastructure, labs, and equipment showed mixed effects, the proportion of female teachers emerged as a consistent determinant of student

achievement, likely reflecting higher educational attainment among female teachers.

The study also documented the adverse impacts of Ecuador's economic downturn and the 2016 earthquake on student performance, particularly for senior-year students. This event underscores the vulnerability of educational outcomes to macroeconomic and environmental shocks, which policymakers must consider in planning resilience strategies for the education sector.

References

- Ali, N., Mukhtar, S., Khan, Y., Ahmad, M., & Khan, Z. U. (2022). Analysis of secondary school students' academic performance and parental involvement in children education at home. *Obrazovanie i Nauka*. <https://doi.org/10.17853/1994-5639-2022-9-118-142>
- Björklund, A., & Salvanes, K. G. (2010). Education and family background: Mechanisms and policies. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.1620398>
- Briones Rugel, F. A., Rivas Requena, J. A., Viteri Narváez, A. E., & Estrada Aguilar, L. M. (2011). La educación en el ecuador, situación y propuesta del sistema de vouchers educativos como alternativa.
- Brunello, G., & Checchi, D. (2005). School quality and family background in italy. *Economics of education review*, *24*(5), 563–577.
- Burtless, G. (2011). *Does money matter?: The effect of school resources on student achievement and adult success*. Brookings Institution Press.
- Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of econometrics*, *18*(1), 5–46.
- Coleman, J. S. (1968). Equality of educational opportunity. *Integrated education*, *6*(5), 19–28.
- Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47. <https://doi.org/10.1257/aer.97.2.31>
- Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of econometrics*, *30*(1-2), 109–126.
- Dookie, S. S. (2014). An evaluation of the levels of parental involvement in activities that support student learning and how it impacts on student performance. *null*. <https://doi.org/null>
- Ferguson, H. B., Bovaird, S., & Mueller, M. P. (2007). The impact of poverty on educational outcomes for children. *Paediatrics & child health*, *12*(8), 701–706.
- Flotts, M. P., Manzi, J., Jiménez, D., Abarzúa, A., Cayuman, C., & García, M. J. (2015). *Informe de resultados terce: Logros de aprendizaje*. UNESCO Publishing.
- Froiland, J. M., Peterson, A., & Davison, M. L. (2013). The long-term effects of early parent involvement and parent expectation in the usa. *School Psychology International*. <https://doi.org/10.1177/0143034312454361>
- Hægeland, T., Raaum, O., & Salvanes, K. G. (2004). *Pupil achievement, school resources and family background* (tech. rep.). Discussion Papers.

- Hanushek, E. A., Rivkin, S. G., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>
- Hoxby, C. M. (2000). The effects of geographic integration and increasing competition in the market for college education. *NBER Working Paper*, 6323, 6323.
- INEVAL. (2017). *Índice socioeconómico, ser estudiante y ser bachiller* (1st).
- Krapohl, E., Rimfeld, K., Shakeshaft, N. G., et al. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences*, 111(42), 15273–15278.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*. <https://doi.org/10.1162/003355399556052>
- Mincer, J. (1974). Schooling, experience, and earnings. *null*. <https://doi.org/10.2307/2521600>
- Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the tennessee class size experiment. *American Educational Research Journal*. <https://doi.org/10.3102/00028312037001123>
- Parveen, S., & Reba, S. H. A. (2016). The impact of parental involvement on children's education. *Humanities and Social Sciences*. <https://doi.org/null>
- Rutkowski, D., & Rutkowski, L. (2013a). Measuring socioeconomic background in pisa: One size might not fit all. *Research in Comparative and International Education*. <https://doi.org/10.2304/rcie.2013.8.3.259>
- Rutkowski, D., & Rutkowski, L. (2013b). Measuring socioeconomic background in pisa: One size might not fit all. *Research in Comparative and International Education*, 8(3), 259–278.
- Shoukat, A., Ilyas, M., Azam, R., & Hussain, A. (2015). Impact of parents' education on children's academic performance. *null*. <https://doi.org/null>
- Stevenson, D. L., & Baker, D. P. (1987). The family-school relation and the child's school performance. *Child Development*. <https://doi.org/10.2307/1130626>
- Toren, N. K. (2013). Multiple dimensions of parental involvement and its links to young adolescent self-evaluation and academic achievement. *Psychology in the Schools*. <https://doi.org/10.1002/pits.21698>
- Verbeek, M. (2017). *A guide to modern econometrics*. John Wiley & Sons.
- Wooldridge, J. M. (2013). Correlated random effects panel data models. *IZA Summer School in Labor Economics*.

- Yunus, K. R. M., & Dahlan, N. A. (2013). Child-rearing practices and socio-economic status: Possible implications for children's educational outcomes. *Procedia - Social and Behavioral Sciences*. <https://doi.org/10.1016/j.sbspro.2013.07.089>
- Zou, W., Anderson, N., Sorin, R., & Hajhashemi, K. (2013). A contextual understanding of mainland chinese parent involvement in their children's primary school years' education. *null*. <https://doi.org/10.2139/ssrn.2637412>