KULTURA I EDUKACJA

**Kultura i Edukacja 2021**
No. 4 (134), pp. 40–70
DOI: 10.15804/kie.2021.04.03
www.kultura-i-edukacja.pl

**Arjan Shahini**[1]

# Inequalities in Albanian Education: Evidence from Large-Scale Assessment Studies

**ABSTRACT**

The study analyzes student, school and district level inequalities of Albanian education system as evidenced in two large-scale assessments. Two main datasets were used for this study, PISA 2018 and the Albanian State Matura Exam 2017. Due to the limited availability of data, the study could only consider a small number of dependent variables at the individual, school, and district level. Utilizing a multilevel analysis, the study observes considerable differences among schools and districts in all three PISA domains and the State Matura Exam. The results were inconclusive regarding shortages of resources at the school and district level. Staff shortage was associated with academic performance in the PISA 2018 dataset, but no statistical association could be identified with the lack of school resources. The analysis of the district financial resources did not show any significant relationship between spending and school performance in the Albanian State Matura Exam. Gender disparities were present in both datasets. Socioeconomic factors, which were measured only in the PISA dataset, had an effect on the student' achievement.

**Keywords:**

PISA, large-scale assessment, Albanian education system, school effectiveness research, Albanian State Matura Exam.

1 Fakultät I, Martin Luther Universität Halle-Wittenberg, Germany.
E-MAIL: arjan.shahini@gmail.com ORCID: 0000-0003-2458-4664

This research was funded and assisted by the Open Society Foundation in Tirana, Albania.

## INTRODUCTION

This paper presents some of the consistent inequalities evidenced in two different large-scale assessments conducted in Albania – PISA 2018, and State Matura Exam (SME) 2017. The purpose of this paper is to look at the factors that might influence students' achievement at the district, school, and individual level. The report aims to analyze whether the inequalities are consistent across both tests, although they measure different outcomes at two different education stages.

The analytical framework of this study is based on the scholarship on school effectiveness research and educational equity to demonstrate how large-scale assessments may be useful tools to monitor inequalities in the Albanian education system and ensure accountability. The main question is whether there are considerable differences between schools and whether they can be attributed to the school characteristics and district-level funding. Based on the empirical observations from the school effectiveness research (see: Chapman et al., 2016), and the latest PISA 2015 analysis (OECD, 2016, p. 184), the main hypothesis is that the unequal distribution of resources in inputs is reflected in students' academic achievement. As pointed out in the PISA 2015 and 2018 reports, not only the insufficient or inadequate education resources but also the unequal distribution of such resources is detrimental to the quality of education provision. The subsequent chapters present the background, the conceptual framework, and the multilevel research design with the PISA 2018 and SME 2017 datasets. The discussion and conclusion chapters focus on the findings and how to improve the use of large-scale assessments as monitoring instruments.

## BACKGROUND

Equity in education has been a long-time concern for the Albanian government. In its most recent national strategy for pre-university education, 2014–2020, the Ministry of Education (2014a) states that one of the strategic guiding principles is to provide "quality and inclusive education" by offering "equal education opportunities" (p. 27). In its assessment of the Albanian policy reforms, UNESCO (2017) evidenced positive changes regarding equity and equality, but also some shortcomings. While Albania has achieved the maximum Gender Parity Index (1.0), according to UNESCO, there are concerns about "income and social class inequality, the rural-urban divide, gender disparities, and discrimination against minorities, and children with disabilities" (p. 42). The government has also set out

a monitoring system and indicators to observe and evaluate the implementation of its strategy. However, most of the indicators cover the collection of information on the expected progress at the input stage (e.g., SCBS5 of the said strategy requires only the reporting PISA results, without any evaluation or feedback mechanism).

In Albania, students are subject to two curriculum-based standardized tests, conducted annually, and two periodical low-stake competence-based large-scale assessments. The two national standardized tests are: (1) "provimi i lirimit" (exit exam – PL) for students graduating from the lower secondary school at the end of compulsory schooling; and (2) "*provimi i maturës shtetërore*" (State Matura Exam – SME), subjecting students graduating from the upper secondary school. While PL is considered a low-stake test (Ministry of Education, 2014a, p. 22), SME is high-stake. Introduced in 2006, SME until recently was also a student selection and academic placement system for the public universities, which enroll about 80 percent of the student population (INSTAT, 2018). The two other large-scale, low-stake and competency-based standardized tests, are (1) PISA, performed since 2002 (a version of PISA 2001 for selected countries), an international large-scale assessment; and (2) the National Assessment for several subjects since 2013 conducted through a representative sample of Albanian primary schools, grade 3 and 5 (National Agency of Examinations, 2014).

The large-scale assessment system has not yet evolved into a methodologically acceptable quality monitoring and accountability system with performance benchmarks and indicators at the school and district level, especially in terms of educational equity. This is recently confirmed by a review of an OECD expert team on the evaluation and assessment system in Albania in 2020 (Maghnouj et al., 2020). Much emphasis has been put to the performance in international assessments, rather than the national curriculum standards. For instance, an intense public debate ensued when the World Bank country report (Gortazar, Kutner, & Inoue, 2014) showed that a large number of students (more than 56 percent) were "functionally illiterate" based on PISA 2012 results. While the considerable and persistent disparities between urban and rural schools have been mostly ignored in both international and national assessments (Schmidt-Neke, 2007).

Average school results of SME have been utilized as indicators of school performance in compiled unofficial school rankings, both in the press and from the local education administrators. Since 2014, Regional Education Directorates (local education administration offices) have been instructed to compile league tables using the school average of SME, the pass rate in SME and the difference of SME results from the school's grades (Ministry of Education, 2014b). This practice has not been followed through consistently since the majority of the dis-

tricts have ceased their online publication of results since 2015. However, league tables based on averages of school performance in such tests are methodically problematic (Goldstein & Thomas, 1996; Meyer, 1993). Moreover, league tables based on performance in standardized tests do not address issues of educational equity. They merely confer responsibility upon the school management and local education authorities, rather than on the funding policies and resource allocation determined by the national and the local government.

## STATE MATURA EXAM

Monitoring the education system and measure inequalities in education through standardized tests is a recent development in Albania. PISA 2001 was the first large-scale assessment conducted in Albania, followed by SME in 2006. The government established the State Matura Exam, a curriculum-based standardized test, to improve the higher education admission system. According to M. Tafaj, advisor to the Prime Minister and later Minister of Education in the years 2009–2013, the SME was conceived in the Democratic Party's master plan of 2004 to replace the universities' admission and entry examination system that was perceived as very corrupt and inefficient. Hence, the SME is not a systemic type of assessment at the national level. Although plagued by allegations of cheating and mismanagement (see, for instance, BalkanWeb, 2015), the overall evaluation of experts and students of SME is positive (IDRA, 2012; OSFA, 2013). Surveys conducted by OSFA (2013) on the student's perception of SME demonstrate that both students and teachers perceive the scoring system as opaque (pp. 14–15), and some doubt its reliability because of the lax oversight during the exam.

SME is both exit and entrance examination. There have been some changes in the number of SME's domains. In 2018 there were five domains: Albanian language (& literature), mathematics, foreign language, and two elective subjects. Albanian language and mathematics are compulsory to graduate from high school (exit examination). Students enroll at their chosen university and study program based on their results (entrance examination) in SME (the admission system is labeled as based on 'merit and preference').

SME is administered by the Educational Services Center (*Qendra e Shërbimeve Arsimore – QSHA*). The test is conducted in two phases, whereby the second phase is a remake of the first exam for those who have not passed in the first phase or from previous years. Scores in SME (from 0–50, maximum 46–50, and 10 points minimum) are assigned to grades (4 to 10) based on a pre-defined distribution (e.g.,

a curve, Vrapi R., Personal Communication, July 6, 2018). The content covered is published, and schools prepare students with different mock exams. The agency issues aggregated annual descriptive reports which are published online. Since its introduction in 2006, there has not been any independent statistical analysis of the student performance of SME results.

## PISA

PISA has been organized by the Organization for Economic Cooperation and Development (OECD) every three years since 2000. The program tests the skills and knowledge of 15-year-old pupils of the OECD and partner countries. Its results enable the evaluation and monitoring of the education system (Lockheed & Wagemaker, 2013). Since 2002, PISA has been part of the monitoring system of Albanian education, especially from an international and comparative outlook. PISA uses a state of the art, transparent, and methodologically adequate assessment and reporting system that is policy-oriented (OECD, 2016; Prenzel, 2012). According to the consortium that administers PISA, the program "assesses the extent to which students near the end of compulsory education have acquired some of the knowledge and skills essential for full participation in society" (OECD, 2016, p. 25). Rather than focusing on the curriculum, like SME, PISA measures the skills and knowledge needed for the future in the domains of reading, mathematical and scientific literacy. For Albania, except for PISA 2000, the PISA survey has been conducted in the domains of reading, mathematics, and scientific literacy. Financial literacy, collaborative problem solving, and other PISA domains have not been measured in Albania.

## SCHOOL CONTEXT VARIABLES

There have been few quantitative studies analyzing the impact of school resources on student performance in Albania. A study by Serra, Barr, and Packard (2011) concluded that "school's physical and human resources were significantly correlated with the proportion of students who excelled in math and Albanian in the 2009 final examination" (p. 7). While in PISA, considering the socioeconomic and demographic background of students, school resources account for about 18 percent of the between-school variance in PISA 2009 reading performance (OECD, 2010, p. 52). PISA 2009 uses a broader definition of school resources, rather than just human and material resources – such as students' learning time

and the intensity of pre-primary education. In PISA 2009, the majority of schools report that they decide on the allocation of the school budget (OECD, 2010, p. 10), and that has a positive effect on the performance, but only for private schools. In 2009, it is reported that a wide variation exists between schools in terms of the extra-curricular opportunities (OECD, 2010, p. 81), which might explain the higher performance of private schools.

The study investigates these variables at the school level: school size, school type, and community type. We expect school size to affect student achievements both in PISA 2015 and in SME. The association between school size and student's cognitive outcomes has been demonstrated in several international and national studies (see the meta-analysis of Scheerens, Hendriks, & Luyten, 2014). The impact of school size depends on whether the school is too large or too small. Studies suggest that there might be an optimal size (e.g., 600 to 900; Andrews, Duncombe, & Yinger, 2002) for which the effect is positive (see: Scheerens et al., 2014). In PISA 2009, larger schools did perform better in reading, but the differences are not statistically significant (Scheerens et al., 2014, p. 30). Larger schools tend to have better human resources, better curricular or extra-curricular opportunities, and that may affect students' performance (Scheerens et al., 2014). Many schools in Albania, especially in rural areas, have poor physical infrastructure. An unpublished study conducted by the Ministry of Education and Science (2011) showed that about a third of rural schools' buildings needed immediate repair or were in a dangerous state. While urban schools are overcrowded, rural schools have small numbers of students (UNESCO, 2017, p. 17).

Another school context variable of relevance, for which we have data, is also the community type. In their meta-analysis on the effects of school context variables, Teddlie, Stringfield, and Reynolds (2000, p. 172) identified a few studies that reported significant findings for the impact of the community type (urban or rural) on the effectiveness of schools. Schools located in the city and urban sectors are expected to be less successful than those of burghs and new town sectors, as shown by Cuttance (as cited in: Teddlie et al., 2000). In another study conducted in the USA, schools in small towns and suburban areas were scoring higher than those in urban and rural areas, results which the authors attributed to the school resources (Teddlie et al., 2000, p. 173). In his meta-analysis, Scheerens (2016, p. 249) found that urban and private schools are more effective then rural and public schools. Researchers also report the effects of interactions between other school characteristics, e.g., school size and community type. Their studies show that urban schools have an advantage in terms of other context variables, such as teacher morale and commitment, resources and school leadership (Teddlie et al., 2000).

In Albania, urban-rural disparities in student performance start in primary education and continue throughout the education system. A study conducted by Serra et al. (2011) with 180 primary schools found that urban schools outperformed rural schools, which also reported higher numbers of school dropouts (p. 5). In the Albanian language exam conducted among the third and fifth grades in 2014, the difference between rural and urban schools was almost half a grade (0.48) in favor of urban schools (QSHA, 2014, p. 31). Disparities continue in the upper secondary as measured both by QSHA (2017b) in the national assessment, again by half a grade difference (0.55, p. 7), and PISA 2012 and 2015 (Lucas et al., 2014; OECD, 2016) , whereby the difference in 2012 was about 50 points in reading, nearly 1.5 years of schooling. The annual reports of SME (see, for example, QSHA, 2017a) show the same considerable differences between rural and urban schools in all domains, in language and reading it was 0.65 grade (2017a, p. 9; there is no information on the statistical analysis conducted, only standard deviations are reported). Disparities are present in other indicators as well. Rural families are disadvantaged in gaining access to quality education (Gjokutaj, 2007, as cited in: UNESCO, 2017). Enrollment rates are lower and public spending is insufficient (World Bank, 2005). Further, the urban-rural disparities have lately moved into the peri-urban areas, reflected in the over-crowded classrooms and the intake from the academically poor equipped students of the rural families (UNESCO, 2017). Urban-rural and public-private inequalities in the SME have been consistent throughout the existence of SME (see Figure 1 and 2 for the last five years).

For the school type, we observe the performance of public *vs.* private schools and the performance of vocational schools. The positive effect of private schools has been evidenced in different international studies (Scheerens et al., 2014). Private schools attract paying students from parents with high socioeconomic status, and they are selective. But the effect of private schools diminishes when student characteristics are accounted for (Lubienski & Lubienski, 2013; Somers, McEwan, & Willms, 2004). A PISA 2012 study found an advantage of private schools only in some countries (Sakellariou, 2017). PISA 2009 for Albania shows that when not accounting for the socioeconomic and demographic background of students and schools, students from private schools perform significantly better than those in public schools (OECD, 2010, p. 44).

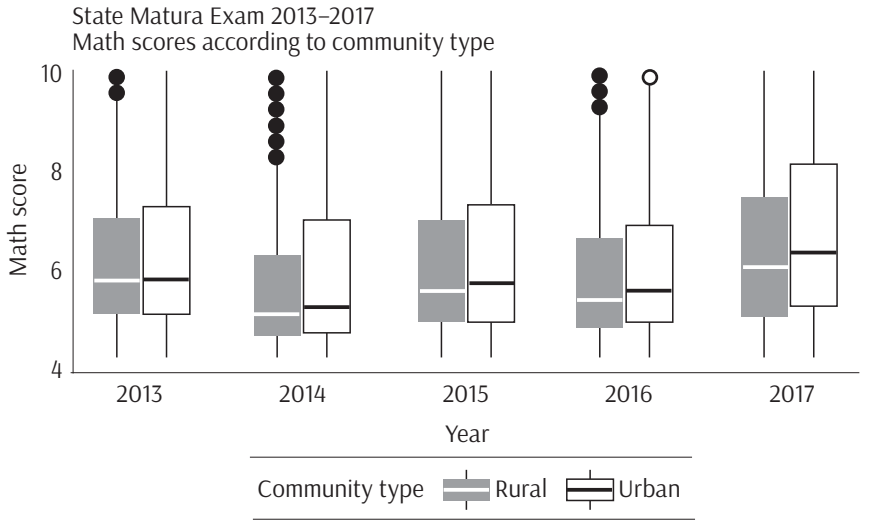


**Figure 1. State Matura Exam 2013–2017: Grade Distribution among Rural and Urban Schools**

Source: QSHA, 2018.

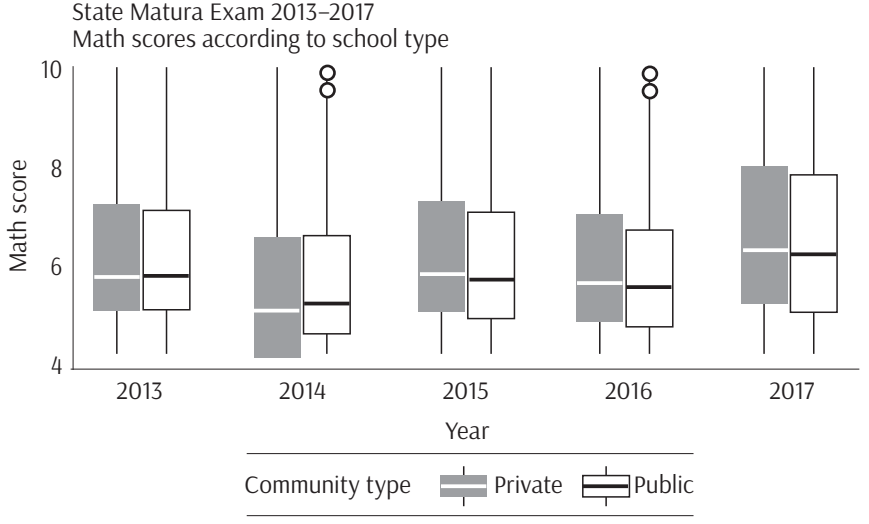


**Figure 2. State Matura Exam 2013–2017: Grade Distribution among Private and Public Schools**

Source: QSHA, 2018.

### DISTRICT LEVEL VARIABLES

Researchers have evidenced an association between variations of the districts in educational resources with the variations of the districts in student outcomes (Fowler & Walberg, 1991; Sander, 1993). Literature on education production function in the USA (e.g., Neely & Diebold, 2016), which looks at the association between input and student outcome, has long found a positive relationship between education spending at district level and achievement, albeit not consistent across studies (Hanushek, 1989). While spending as input may also be considered an index of quality of education (Rizzuto & Wachtel, 1980), the socioeconomic status of the districts also has an effect on the learning outcomes (Fowler & Walberg, 1991). In this study, we will only focus on the spending per student in education at the district level.

### INDIVIDUAL-LEVEL VARIABLES

While the PISA 2018 dataset contains some relevant student-level data, the Albania SME dataset only contains gender. The variables under analysis in the PISA 2018 model are gender and socioeconomic status.

Various studies such as PIRLS, TIMSS, PISA, indicate that there is a gender gap in educational outcomes across many countries in the developed world (Borgonovi, Ferrara, & Maghnouj, 2018). The gender disparities vary according to the subject and that may be attributed to the family (Jacobs & Bleeker, 2004) or the school environment (Figlio et al., 2016). For example, in Italy, girls underperform in mathematic achievement and other science-related fields (Contini, Di Tommaso, & Mendolia, 2017). In the United States, the very same gender differences in favor of the boys are evidenced since early childhood (Penner & Paret, 2008). We expect the same results in our study concerning gender disparities.

We also expect a positive relationship between student achievement and their socioeconomic background. This will be analyzed in the PISA 2018 dataset. Studies reviewed by Sirin (2005) suggest that the socioeconomic background of students does affect student outcome, although this effect depends on the school environment as well, while the effect sizes vary between countries (Thomson, 2018). For instance, the PISA 2003 and PISA 2015 studies conducted in OECD and some developing countries indicate that there is an association between socioeconomic status and test performance (OECD, 2016). In China, Liu, Peng, and Luo (2020) confirmed in their meta-analysis only a small correlation between socioeconomic status and academic achievement.

## METHODOLOGY

The availability of the large datasets, the quantitative approach in the data collection method (PISA 2018), the aim of research to generalize the findings, and the research question place this study under the quantitative tradition. This study is a multilevel analysis because both the datasets (PISA 2015 and SME) and the model are hierarchical in their structure (see: Goldstein & Thomas, 1996; Gelman & Hill, 2006). Students are nested in schools and schools nested in districts. Hierarchical linear models, or multilevel modeling, takes the nested structure of the data into account by estimating the effects at each level (Raudenbush & Willms, 2014). For the PISA 2018 dataset, we conducted a two-level analytic model (student-school). For the SME, we run a three-level analytical model (student-school-district).

### RESEARCH QUESTIONS

The research questions in this study are:

1. Are there any differences in student performance between schools and districts? Are there any similar results to be expected in the PISA and the SME dataset?
2. Are there any differences in the type of schools or community?
3. May these differences be attributed to the shortages of school resources?
4. May these differences be attributed to the unequal distribution of resources at the district level, as measured by spending per student?
5. Are there any gender differences?
6. Are there any differences concerning the socioeconomic status of students in the PISA dataset?

While for the first and second questions we used the PISA and SME datasets, in the third we have used PISA 2018, and for the last question SME 2017.

### STUDY PARTICIPANTS

The PISA 2018 sample contains 6359 individual observations nested in 327 schools. The SME 2017 dataset contains the whole population of high school graduates of 32,976 students, 414 schools nested in 57 municipalities.

### INDEPENDENT VARIABLES

In the PISA survey, the data on the school resources have been obtained through the questionnaire of the school principals. Principals had to answer questions con-

cerning the lack of infrastructure, the quality, and shortage of staff, etc. At the student's level, we can control for a set of students' characteristics that have been measured by PISA. In this study, we have included gender, the socioeconomic background indexes compiled by PISA (ESCS, CULTPOSS, HEDRES). At Level 2, i.e., the school level, we have included these aggregated variables in the model: (i) the school socioeconomic intake (the school average of the student ESCS index); (ii) education shortages, compiled in an index by PISA, (EDUSHORT); (iii) shortage of staff, an index (STAFFSHORT) that was also derived from four questions in the principal's survey; (iv) school type (private/public); (v) community type (urban/rural).

For the analysis with the State Matura Exam 2017 dataset, the study focuses on the school and district level variables. At the school level, the dataset contains data on the school characteristics such as the school type (private/public), community type (urban/rural), and school size. District level variables in this study comprise financial resources for education (expressed in expenditure per pupil).

#### DEPENDENT VARIABLES

The dependent variables for the PISA 2018 dataset are student's performance in science, literacy, and mathematics. PISA score has a scale with a mean of 500 and a standard deviation of 100 (continuous). While in the multivariate model we use all plausible values, in the multilevel model, the study uses the averages of the ten plausible values in science, math, and literacy for the PISA 2018 dataset (similar methods were applied by: Laukaityte & Wiberg, 2017; Sakellariou, 2017). Multivariate regression was conducted for the domain of reading with the statistical software R and the IEA IDB Analyser (2019). Lme4 package in R (Bates et al., 2015) and HLM for Windows (Raudenbush, Bryk, & Congdon, 2013) were used for the multilevel analysis. Previous studies (see: Laukaityte & Wiberg, 2017) have shown that the results differ on whether all ten plausible values or the average of all plausible values are used in the analyses. Using all plausible values is very compute- and time-intensive. In the multilevel model only, a single value was used, the average of the plausible values. There are considerable differences in the intercept and the standard errors of the coefficients between using all plausible values or only one single value. However, the association between the variables and the effect size does not change.

The dependent variables for the SME dataset are student's average scores from the tests of math, literature, and two selected subjects for the year 2017. The SME aggregated data was provided by the Educational Services Center in Albania (QSHA).

**Table 1. List of Variables Used in the Models**

| Variable name | Description |
|---|---|
| ***PISA Models* – Dependent Variables** | |
| MATH | Student's mathematics score, continuous variable |
| SCIE | Students score in the domain of science, continuous variable |
| READ | Students score in reading, continuous variable |
| PVRCUN | Cognitive Process Subscale of Reading – Understanding |
| PVRCLI | Cognitive Process Subscale of Reading – Locate information |
| PVRCER | Cognitive Process Subscale of Reading – Evaluate and Reflect |
| **Independent Variables – *Student Level*** | |
| Female | Students' gender, the control group is female |
| ESCS | Index of socioeconomic status (compiled by PISA) |
| HEDRES | Home educational resources (Index, PISA) |
| CULTPOSS | Cultural possessions at home (Index, PISA) |
| WEALTH | Family wealth |
| **Independent Variables – *School Level*** | |
| ESCS_M | School's socioeconomic intake, continuous variable<br>Aggregated individual ESCS index at the school level |
| ESCS_SD | School's diversity of socioeconomic intake, continuous variable (the standard deviation of ESCS index) |
| EDUSHORT | Education shortage, index, continuous variable |
| STAFFSHORT | Staff shortage, index, continuous variable |
| PROAT5AM | The proportion of teachers with masters, continuous variable |
| CLSIZE | Class size, continuous variable |
| Rural | Rural schools, dummy variable (0 for urban, 1 for rural) |
| PRIVATESCH | School type, dummy variable (0 for private, 1 for public) |
| Community | Location of the school in a village, small town, town, or city |
| ***State Matura Exam* – Dependent Variables** | |
| GRADE | Student's performance in State Matura Exam, continuous variable (from 4 to 10) |
| **Independent Variables – *School Level*** | |
| LgSchool | School size, dummy variable (0 = other, 1 = large school) |
| SmSchool | School size, dummy variable (0 = other, 1 = small school)<br>*Medium size schools as controlling variable* |
| Rural | Rural schools, dummy variable (0 for urban, 1 for rural) |
| Public | School type, dummy variable (0 for private, 1 for public) |
| Geo | School's geographical location (North, Middle, South Albania); categorical variable, with reference group South |

| Variable name | Description |
|---|---|
| Profile | School's profile: gymnasium, vocational, part-time (categorical variable, the reference category is gymnasium) |
| **Independent Variable – *District Level*** | |
| Log(CsP2017) | Log of spending per capita for the secondary level of education at the district level for the year 2017 (continuous variable). The data is for the municipalities |

**Table 2. Summary Table of the Main Variables**

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|
| **Student Level – PISA 2018** | | | | | | | |
| CNTSTID | 6,359 | | | | | | |
| Gender [Female] | 6,359 | 1.50 | 0.50 | 1 | 1 | 2 | 2 |
| MATH | 6,359 | 437.73 | 73.31 | 169.81 | 384.78 | 487.19 | 696.45 |
| SCIE | 6,359 | 417.61 | 68.16 | 197.20 | 368.93 | 464.59 | 664.02 |
| READ | 6,359 | 340.89 | 64.23 | 151.56 | 295.23 | 384.42 | 559.96 |
| PVRCLI | 6,359 | 395.53 | 79.95 | 142.01 | 338.58 | 451.35 | 653.88 |
| PVRCUN | 6,359 | 405.04 | 81.99 | 163.66 | 346.47 | 462.04 | 679.09 |
| PVRCER | 6,359 | 404.85 | 85.26 | 120.86 | 342.97 | 461.22 | 702.25 |
| ESCS | 6,277 | -0.83 | 0.96 | -5.34 | -1.56 | -0.13 | 3.22 |
| CULTPOSS | 6,126 | -0.39 | 0.72 | -1.82 | -0.88 | 0.07 | 1.95 |
| HEDRES | 6,162 | -0.50 | 1.04 | -4.41 | -1.30 | 0.05 | 1.21 |
| WEALTH | 6,241 | -1.23 | 1.07 | -7.55 | -1.89 | -0.56 | 4.48 |
| PRIVATESCH | 6,359 | 1.89 | 0.312 | 1 | 1 | 2 | 2 |
| CNTSTUID | 6,359 | | | | | | |
| **School Level Variables – PISA 2018** | | | | | | | |
| CNTSCHID | 327 | | | | | | |
| EDUSHORT | 323 | 0.40 | 1.08 | -1.42 | -0.31 | 1.25 | 2.96 |
| STAFFSHORT | 324 | -0.62 | 1.01 | -1.47 | -1.46 | -0.01 | 4.04 |
| PROAT5AM | 317 | 0.62 | 0.38 | 0.00 | 0.20 | 0.98 | 1.00 |
| SCHSIZE | 126 | 421.79 | 382.98 | 1.00 | 132.25 | 577.25 | 1,870.00 |
| CLSIZE | 325 | 22.95 | 7.52 | 13.00 | 18.00 | 28.00 | 53.00 |

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|
| **State Matura Exam – 2017** | | | | | | | |
| StudentID | 32,976 | | | | | | |
| SchoolID | 414 | | | | | | |
| Municipality | 57 | | | | | | |
| District | 37 | | | | | | |
| Grade | 32,976 | 6.90 | 1.303 | 4.000 | 5.836 | 7.914 | 9.940 |
| SCHOOLSIZE | 32,976 | 2.04 | 0.89 | 1 | 1 | 3 | 3 |
| Gender[Female] | 32,976 | 0.51 | 0.5 | 0 | 0 | 1 | 1 |
| Profile (Gymansium, Vocational, Part-Time) | 32,976 | 3414 | 3384 | 1 | 1 | 8 | 9 |
| Rural | 32,976 | 0.328 | 0.469 | 0 | 0 | 1 | 1 |
| Urban | 32,976 | 0.672 | 0.469 | 0 | 0 | 1 | 1 |
| Private | 32,976 | 0.123 | 0.329 | 0 | 0 | 0 | 1 |
| Public | 32,976 | 0.877 | 0.329 | 0 | 1 | 1 | 1 |
| CsP2017 | 32,946 | 288,012 | 250,344 | 1,325 | 55,123 | 495,755 | 715,459 |

**STATISTICAL MODELS**

The first model is a simple multivariate linear regression with the individual-level variables of gender, socioeconomic indicators (Index of economic, social, and cultural status), and some school-level variables as factors. The multivariate linear model is compared to the multilevel model. Then, following the recommendations of Hox, Moerbeek, and Van de Schoot (2010) and Lecki (2010), the first analytical model is empty (random intercept model), with no independent variables. The first model (equation 1) shows the decomposed variance of the dependent variable at the student (within school) and school level (between schools). Statistically significant differences between schools would demonstrate that schools differ in terms of average student achievement. The portion of variance is calculated through the ICC (intraclass correlation), which may also be interpreted as correlations of student achievement within schools.

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \qquad (1)$$

In the second model, both in PISA and SME, we added the independent variables. For the second model, we have used the municipalities instead of school districts because the variable "spending per student at secondary level" was available at this level.

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + u_{0j} + e_{ij.} \qquad (2)$$

The statistical analysis was conducted primarily with R (R Development Core Team, 2014), package *lme4* (Bates et al., 2015). Packages *sjPlot* (Lüdecke, 2018a) and *sjstats* (Lüdecke, 2018b) in R were used for the significance tests and the plots.

## RESULTS

### MULTIVARIATE ANALYSIS

We firstly looked at the influence of individual-level and school-level variables through a multivariate regression model for all the domains. In the model presented (see: Figure 3, for details see: Table 3), we have used all ten plausible values in SPSS (the results are similar to the results provided with the software of the International Association for the Evaluation of Educational Achievement, IEA

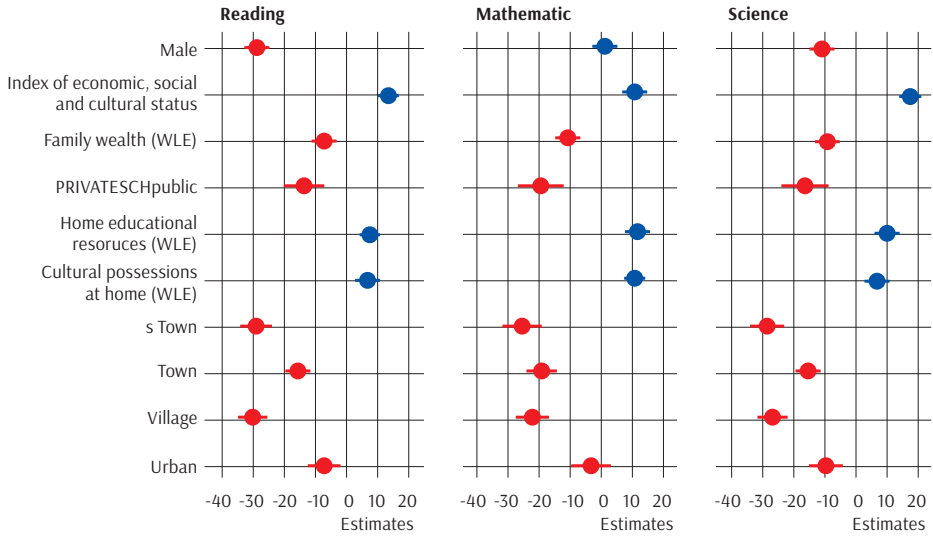


**Figure 3. Plots of Multivariate Regressions for All PISA 2018 Domains**

**Table 3. Multivariate Regression Model for Reading, Mathematics, and Science**

| | READ | | | | MATH | | | | SCIE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Predictors* | *Esti-mates* | *std. Error* | *CI* | *p* | *Esti-mates* | *std. Error* | *CI* | *p* | *Esti-mates* | *std. Error* | *CI* | *p* |
| (Intercept) | 394.36 | 2.41 | 389.64 – 399.08 | **<0.001** | 476.68 | 2.93 | 470.94 – 482.42 | **<0.001** | 469.14 | 2.64 | 463.96 – 474.32 | **<0.001** |
| **Female: Male** | -26.29 | 1.47 | -29.17 – -23.41 | **<0.001** | 0.80 | 1.79 | -2.70 – 4.30 | 0.655 | -11.63 | 1.61 | -14.79 – -8.47 | **<0.001** |
| Index of economic, social and cultural status | 12.74 | 1.22 | 10.36 – 15.13 | **<0.001** | 9.38 | 1.48 | 6.48 – 12.28 | **<0.001** | 15.09 | 1.34 | 12.47 – 17.71 | **<0.001** |
| Family wealth(WLE) | -6.54 | 0.99 | -8.47 – -4.61 | **<0.001** | -10.38 | 1.20 | -12.73 – -8.03 | **<0.001** | -9.98 | 1.08 | -12.11 – -7.86 | **<0.001** |
| PRIVATESCHpublic | -12.42 | 2.95 | -18.21 – -6.63 | **<0.001** | -19.05 | 3.59 | -26.10 – -12.01 | **<0.001** | -16.73 | 3.24 | -23.08 – -10.37 | **<0.001** |
| Home educational resources(WLE) | 7.31 | 0.91 | 5.52 – 9.10 | **<0.001** | 10.38 | 1.11 | 8.21 – 12.55 | **<0.001** | 7.90 | 1.00 | 5.94 – 9.86 | **<0.001** |
| Cultural possessions at home(WLE) | 6.38 | 1.18 | 4.07 – 8.69 | **<0.001** | 8.34 | 1.43 | 5.54 – 11.15 | **<0.001** | 4.67 | 1.29 | 2.14 – 7.21 | **<0.001** |
| **Community: s Town** | -26.25 | 2.38 | -30.93 – -21.58 | **<0.001** | -24.34 | 2.90 | -30.02 – -18.66 | **<0.001** | -27.31 | 2.62 | -32.44 – -22.18 | **<0.001** |
| Community: Town | -13.80 | 1.93 | -17.59 – -10.01 | **<0.001** | -18.30 | 2.35 | -22.91 – -13.69 | **<0.001** | -15.38 | 2.12 | -19.54 – -11.23 | **<0.001** |
| **Community: Village** | -27.70 | 2.18 | -31.97 – -23.43 | **<0.001** | -21.36 | 2.65 | -26.56 – -16.17 | **<0.001** | -26.18 | 2.39 | -30.86 – -21.49 | **<0.001** |
| Location: Urban | -6.70 | 2.40 | -11.41 – -1.99 | **0.005** | -3.25 | 2.92 | -8.98 – 2.48 | 0.266 | -9.72 | 2.64 | -14.89 – -4.55 | **<0.001** |
| **Observations** | 6079 | | | | 6079 | | | | 6079 | | | |
| $R^2$ / $R^2$ adjusted | 0.211 / 0.209 | | | | 0.106 / 0.105 | | | | 0.157 / 0.156 | | | |

Database Analyser). We have observed from the model that for every unit increase in the socioeconomic status of students there is an increase from 9.5 to 15.5 points in all PISA domains. A significant performance gap exists between males and females in reading (-32 points) and science (-4.6 points). In mathematics, there are no significant gender differences. In comparison to students from the cities, those in small towns and rural areas are expected to perform poorer. The model also shows that students from the villages perform better than those in the urban areas, but that could be attributed to the fact that many students from the villages (about 33 percent) visit a school in the city. Surprisingly, the relationship between family wealth and performance is negative in all three domains. As expected, based on the previous PISA cycles, being in a private school is positively associated with the performance in all three domains. The model shows that students from private schools are expected to perform better than those in public schools.

Since it constitutes the focus of the survey for this PISA cycle, we have also designed a model to look at the reading performance more in detail (see: Figure 4, for more details see: Table 4). Student performance in reading demonstrates a considerable gap between males and females. Based on the model, it is expected that males have 26.5 points less than females. The gender differences in favor of females are consistent in all the reading subscales, whereby the largest is in the cognitive processes of understanding information and evaluating and reflecting. The cognitive assessment of these subscales comprises three main aspects of the reading typology in text processing: information location, understanding, and evaluation (OECD, 2019a). The results of the multivariate regression model show that males perform poorly in comparison to girls in understanding and evaluating the text for its credibility. Socioeconomic status, cultural possessions, and educational resources at home have a positive association with reading performance and its subscales. Considering previous research (Sirin, 2005; Thomson, 2018), it is not a surprise that cultural possession and home education resources are better predictors for a positive score. An increase of one unit in the socioeconomic status index of the students is associated with an increase of 13 points in their performance in reading. The findings are consistent with PISA 2018 report for Albania (OECD, 2019b, p. 4), which states that students from an advantaged socioeconomic background outperform those from disadvantaged by 61 score points in PISA 2018. Class size does not have a significant effect on the reading performance at all subscales. Students from the schools in small towns, villages, and urban areas perform poorly, especially in comparison to other students in the cities. The differences are much sharper for schools located in the villages, as the model shows. The multilinear model explains about 20 percent of the variability of the response variable, the reading fluency in this case.

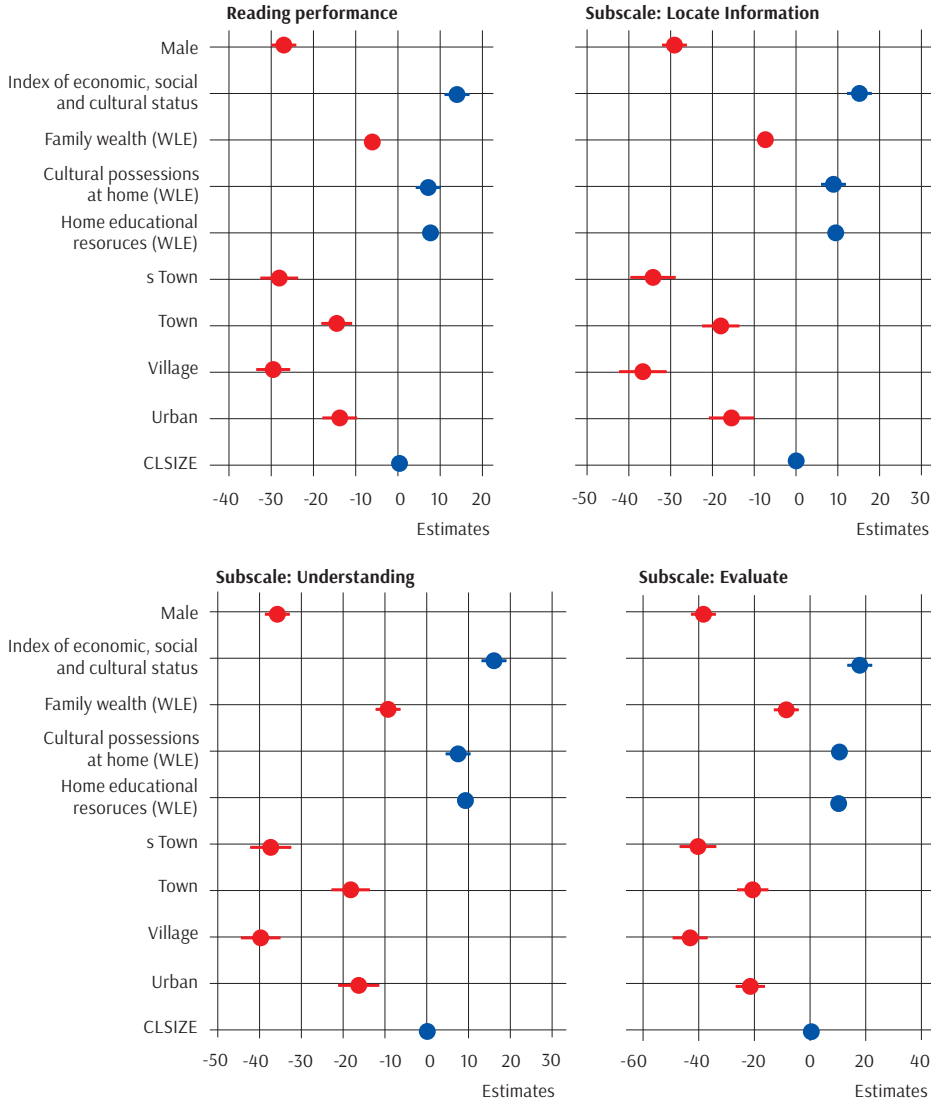


**Figure 4. Plots of Regression Models for Reading Performance and Its Three Subscales, PISA 2018**

**Table 4. Multivariate Regression Coefficients for Students' Performance in Reading Fluency, PISA 2018**

| | READ | | ■ Locate Information | | ■ Understand | | ■ Evaluate and Reflect | |
|---|---|---|---|---|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *Estimates* | *CI* | *Estimates* | *CI* | *Estimates* | *CI* |
| (Intercept) | 386.04 | 379.17 – 392.90 | 454.85 | 446.17 – 463.52 | 464.47 | 455.59 – 473.35 | 467.94 | 458.81 – 477.07 |
| Female [Male] | -26.55 | -29.44 – -23.67 | -29.38 | -33.02 – -25.74 | -36.13 | -39.83 – -32.44 | -35.47 | -39.30 – -31.63 |
| ESCS | 13.13 | 10.74 – 15.51 | 16.10 | 13.09 – 19.12 | 16.45 | 13.43 – 19.48 | 16.90 | 13.73 – 20.07 |
| WEALTH | -6.27 | -8.20 – -4.33 | -7.29 | -9.74 – -4.84 | -8.92 | -11.42 – -6.42 | -8.46 | -11.03 – -5.88 |
| CULTPOSS | 6.42 | 4.11 – 8.73 | 8.19 | 5.26 – 11.11 | 8.00 | 5.04 – 10.97 | 9.04 | 5.96 – 12.11 |
| HEDRES | 7.11 | 5.31 – 8.90 | 9.37 | 7.11 – 11.64 | 10.09 | 7.80 – 12.39 | 9.55 | 7.16 – 11.93 |
| Community [sTown] | -27.86 | -32.47 – -23.25 | -33.59 | -39.42 – -27.77 | -38.29 | -44.02 – -32.57 | -37.75 | -43.89 – -31.62 |
| Community [Town] | -14.62 | -18.40 – -10.85 | -18.75 | -23.52 – -13.98 | -18.19 | -22.84 – -13.54 | -19.15 | -24.17 – -14.13 |
| Community [Village] | -29.30 | -33.60 – -24.99 | -37.03 | -42.48 – -31.59 | -40.38 | -45.96 – -34.81 | -40.35 | -46.09 – -34.62 |
| Location [Urban] | -13.89 | -18.00 – -9.78 | -17.17 | -22.36 – -11.98 | -16.59 | -21.94 – -11.24 | -21.77 | -27.23 – -16.30 |
| CLSIZE | 0.18 | -0.06 – 0.41 | 0.09 | -0.21 – 0.39 | 0.18 | -0.12 – 0.48 | 0.26 | -0.05 – 0.57 |
| Observations | 6076 | | 6076 | | 6076 | | 6076 | |
| $R^2$ / $R^2$ adjusted | 0.209 / 0.207 | | 0.199 / 0.197 | | 0.211 / 0.210 | | 0.212 / 0.211 | |

## MULTILEVEL ANALYSIS

The second research question looks at the differences between schools. We also checked whether a multilevel analysis is a better fit than a simple OLS regression (Lunn et al., 2000), and the multilevel model is a much better method.

The results demonstrate that there are considerable differences between schools, both in the PISA and the SME dataset. In PISA 2018, a two-level multilevel analysis shows that 27% of the variance in reading score is attributable to differences between schools, and 24% and 26% respectively in mathematics and science (see: Table 5). The variance between schools, or the ICC (*intraclass correlation*), could also be interpreted as the correlation among students within schools.

**Table 5. Intercept Only for the PISA 2018 Dataset**

| *Predictors* | **READ** *(CI estimates)* | **MATH** *(CI estimates)* | **SCIE** *(CI estimates)* |
|---|---|---|---|
| (Intercept) | 334.77 (330.70 – 338.84) | 435.29 (430.84 – 439.75) | 412.72 (408.43 – 417.00) |
| **Random Effects** | | | |
| $\sigma^2$ | 2946.84 | 3983.63 | 3333.71 |
| $\tau_{00}$ | $1080.67_{\text{CNTSCHID}}$ | $1256.20_{\text{CNTSCHID}}$ | $1190.86_{\text{CNTSCHID}}$ |
| ICC | 0.27 | 0.24 | 0.26 |
| N | $327_{\text{CNTSCHID}}$ | $327_{\text{CNTSCHID}}$ | $327_{\text{CNTSCHID}}$ |
| Observations | 6359 | 6359 | 6359 |
| Marginal $R^2$ / Conditional $R^2$ | 0.000 / 0.268 | 0.000 / 0.240 | 0.000 / 0.263 |

In the SME model, the school differences "explain" between 30% and 35% of the student's performance variance, and between 5% to 18% of the variance is found at the district level (see: Table 6). The school variance is relatively high for the SME, and it confirms the findings in other developing countries. Concerning school differences, we conclude that schools do have an impact on student's learning outcomes, as measured by both an international and a national large-scale assessment test. Effect sizes, measured by comparing the average scores of the most effective with the least effective schools (Scheerens et al., 2014), are larger than 0.8 for the SME (for 2017 Cohen's $d = 9.95$, effect size $r = 0.97$).

**Table 6. Intercept Only for the SME Datasets**

| | 2013 (Std. Error) | 2014 | 2015 | 2016 | 2017 |
|---|---|---|---|---|---|
| Fixed part | | | | | |
| Intercept | 6.42 (0.07) | 5.87 (0.06) | 6.42 (0.1) | 6.56 (0.06) | 6.5 (0.06) |
| Random part | | | | | |
| District level | 0.12 | 0.08 | 0.06 | 0.06 | 0.08 |
| School level | 0.51 | 0.51 | 0.53 | 0.48 | 0.55 |
| Pupil level | 0.92 | 1.16 | 1.09 | 0.88 | 1.01 |
| Variance partition | | | | | |
| School level | 33.1% | 29.5% | 31.8% | 33.5% | 33.4% |
| District level | 7.6% | 4.8% | 3.6% | 4.6% | 4.7% |
| Number of students | 38,102 | 37,652 | 35,911 | 34,655 | 32,347 |
| Number of schools | 433 | 439 | 439 | 440 | 497 |
| Number of districts | 38 | 38 | 38 | 38 | 38 |

To study the effect of the school resources, we concentrated exclusively on the PISA dataset. We included the EDUSHORT and STAFFSHORT as variables in the model; indexes compiled by the principals' survey (as the number of schools, n) to look at the effect of school resources and the shortage of staff as an indicator of lack of resources. In the PISA survey, some 40 percent of students were in schools, where it was reported that the lack of physical infrastructure was detrimental to the school's capacity to provide instruction. According to the OECD (2019), about 8 percent of students attended an advantaged school, while school principals perceived the shortage of staff was an obstacle to instruction.

From the multilevel model in our study (see: Figure 5), we can confirm the negative association between the lack of staff and students' performance in PISA. For every increase in one unit of the STAFFSHORT variable, there is a decrease from four to five points in the PISA score, in all the domains. Also, the model shows that the effect of private schools, which, as expected, is positive, is not significant in the domain of reading. While the socioeconomic characteristics of schools, their socioeconomic student intake, is significant in all the domains. The inadequacy and insufficiency of resources are negatively associated with student performance (shown in Table 7), but the variables measuring the shortage and the percentage of teachers with master's degrees are not statistically significant. In the reading component, the effects of staff shortage and the location of schools in small towns and villages are also negative.

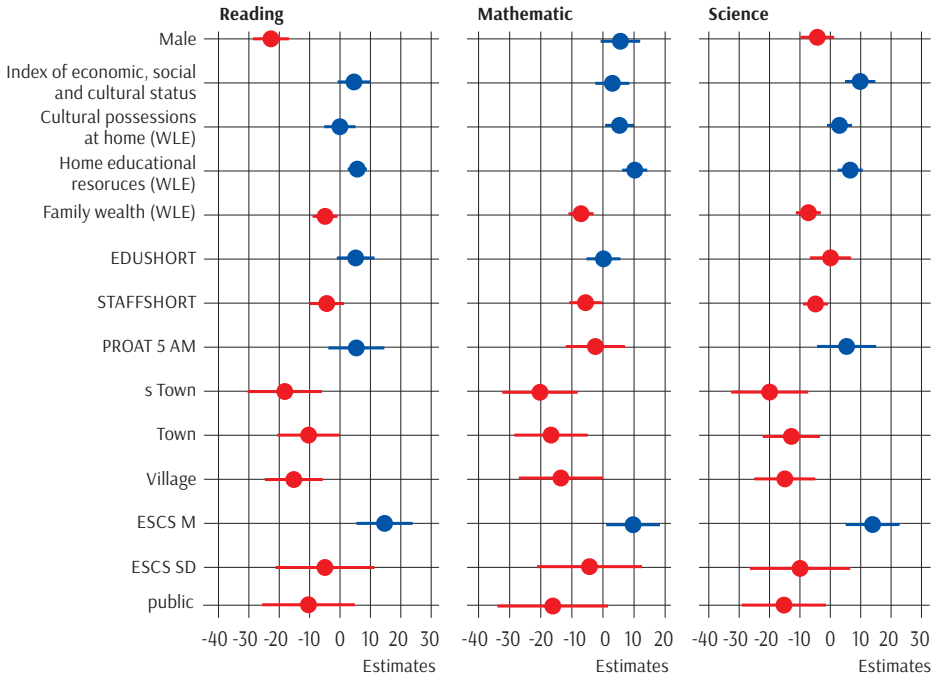


**Figure 5. Regression Estimates for Literacy, Mathematics, and Science (PISA 2018)**

State Matura Exam grades also show significant associations between school type and community type for urban and private schools (for details, see: Figure 6 and Table 8). Keeping all other variables constant, we can predict that a student coming from a public school will have lower grades, and a further decrease in performance if coming from a rural school or a small school. Vocational schools also have a significant negative relationship with student performance, which could be attributed to the fact that these schools attract students that are relatively academically ill-prepared. The education budget at the district level does not show any significant effect, as the region of the school whether the school is in central or southern Albania.

At the individual level, the PISA 2018 dataset suggests that like the multivariate model gender and students' socioeconomic status have a positive effect, but strangely enough, wealth seems to have a negative effect in this model (see: Figure 5). The same gender differences in favor of females are observed in the State Matura Exam as well.

**Table 7. Model for the PISA 2018 Dataset**

| | READ | | | | MATH | | | | SCIE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Predictors* | *Esti-mates* | *std. Error* | *CI* | *p* | *Esti-mates* | *std. Error* | *CI* | *p* | *Esti-mates* | *std. Error* | *CI* | *p* |
| (Intercept) | 384.63 | 7.94 | 369.08 – 400.19 | **<0.001** | 470.09 | 9.89 | 450.70 – 489.48 | **<0.001** | 461.50 | 8.76 | 444.34 – 478.66 | **<0.001** |
| **Female: Male** | -21.84 | 1.41 | -24.60 – -19.08 | **<0.001** | 4.93 | 1.71 | 1.58 – 8.27 | **0.004** | -6.66 | 1.55 | -9.70 – -3.62 | **<0.001** |
| Index of economic, social and cultural status | 6.83 | 1.19 | 4.50 – 9.16 | **<0.001** | 3.80 | 1.44 | 0.97 – 6.63 | **0.008** | 9.57 | 1.31 | 7.00 – 12.14 | **<0.001** |
| Cultural possessions at home(WLE) | 5.48 | 1.11 | 3.31 – 7.65 | **<0.001** | 7.15 | 1.34 | 4.52 – 9.78 | **<0.001** | 3.50 | 1.22 | 1.11 – 5.89 | **0.004** |
| Home educational resources(WLE) | 7.42 | 0.86 | 5.72 – 9.11 | **<0.001** | 10.33 | 1.05 | 8.28 – 12.38 | **<0.001** | 7.67 | 0.95 | 5.81 – 9.54 | **<0.001** |
| Family wealth(WLE) | -5.33 | 0.93 | -7.15 – -3.50 | **<0.001** | -8.67 | 1.13 | -10.89 – -6.46 | **<0.001** | -8.37 | 1.03 | -10.38 – -6.36 | **<0.001** |
| EDUSHORT | 1.28 | 1.89 | -2.43 – 4.99 | 0.499 | 0.52 | 2.36 | -4.12 – 5.15 | 0.827 | -0.38 | 2.09 | -4.47 – 3.72 | 0.857 |
| **STAFFSHORT** | -4.29 | 1.83 | -7.87 – -0.70 | **0.019** | -5.71 | 2.28 | -10.18 – -1.23 | **0.012** | -4.28 | 2.02 | -8.24 – -0.33 | **0.034** |
| PROAT 5 AM | 4.19 | 4.20 | -4.05 – 12.42 | 0.319 | -1.46 | 5.25 | -11.75 – 8.83 | 0.781 | 3.59 | 4.64 | -5.49 – 12.68 | 0.438 |
| **Community: s Town** | -18.99 | 5.59 | -29.94 – -8.04 | **0.001** | -19.50 | 6.99 | -33.20 – -5.81 | **0.005** | -19.54 | 6.17 | -31.63 – -7.46 | **0.002** |
| Community: Town | -9.68 | 4.72 | -18.93 – -0.43 | **0.040** | -14.02 | 5.91 | -25.61 – -2.44 | **0.018** | -11.52 | 5.21 | -21.73 – -1.31 | **0.027** |
| **Community: Village** | -15.14 | 4.96 | -24.86 – -5.41 | **0.002** | -12.26 | 6.20 | -24.41 – -0.12 | **0.048** | -14.51 | 5.48 | -25.25 – -3.77 | **0.008** |
| ESCS M | 15.83 | 3.76 | 8.46 – 23.20 | **<0.001** | 9.46 | 4.68 | 0.27 – 18.64 | **0.044** | 12.60 | 4.15 | 4.47 – 20.73 | **0.002** |
| **ESCS SD** | -4.14 | 7.61 | -19.06 – 10.78 | 0.586 | -2.13 | 9.47 | -20.69 – 16.44 | 0.822 | -10.38 | 8.40 | -26.84 – 6.08 | 0.216 |
| PRIVATESCH: public | -10.24 | 6.24 | -22.48 – 1.99 | 0.101 | -17.95 | 7.80 | -33.25 – -2.66 | **0.021** | -15.88 | 6.89 | -29.38 – -2.38 | **0.021** |
| **Random Effects** | | | | | | | | | | | | |
| $\sigma^2$ | 11120.12 | | | | 16321.48 | | | | 13497.53 | | | |
| $\tau_{00}$ | $528.44_{CNTSCHID}$ | | | | $840.46_{CNTSCHID}$ | | | | $644.28_{CNTSCHID}$ | | | |
| ICC | 0.05 | | | | 0.05 | | | | 0.05 | | | |
| N | $295_{CNTSCHID}$ | | | | $295_{CNTSCHID}$ | | | | $295_{CNTSCHID}$ | | | |
| Observations | 5933 | | | | 5933 | | | | 5933 | | | |
| Marginal $R^2$ / Conditional $R^2$ | 0.057 / 0.100 | | | | 0.027 / 0.074 | | | | 0.039 / 0.083 | | | |

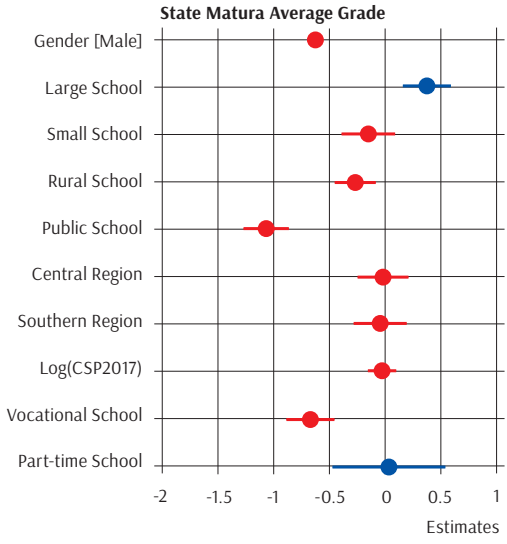


**Figure 6. Regression Model for State Matura Exam, 2017**

**Table 8. Hierarchical Level Model for Student Achievement, SME 2017**

| | Model 0 - GRADE | | | | Model 1 - GRADE | | | |
|---|---|---|---|---|---|---|---|---|
| *Predictors* | *Esti-mates* | *Std. Error* | *CI* | *p* | *Esti-mates* | *Std. Error* | *CI* | *p* |
| *Fixed effects* | | | | | | | | |
| **(Intercept)** | 6.56 | 0.06 | 6.44 – 6.67 | **<0.001** | 8.36 | 0.59 | 7.59 – 9.13 | **<0.001** |
| Student level | | | | | | | | |
| **Gender [Male]** | | | | | -0.65 | 0.01 | -0.67 – -0.63 | **<0.001** |
| School level | | | | | | | | |
| **Large School** | | | | | 0.34 | 0.10 | 0.13 – 0.54 | 0.001 |
| Small School | | | | | -0.24 | 0.08 | -0.40 – -0.09 | 0.002 |
| **Vocational School** | | | | | -0.67 | 0.11 | -0.89 – -0.46 | <0.001 |
| Part-time School | | | | | 0.03 | 0.29 | -0.53 – 0.59 | 0.905 |
| **Rural School** | | | | | -0.31 | 0.08 | -0.47 – -0.16 | <0.001 |
| Public school | | | | | -1.06 | 0.09 | -1.24 – -0.89 | <0.001 |
| **District level** | | | | | | | | |
| Central region | | | | | -0.02 | 0.10 | -0.22 – 0.17 | 0.804 |
| **Southern region** | | | | | -0.04 | 0.10 | -0.24 – 0.16 | 0.701 |
| Log(CsP2017) [Budget] | | | | | -0.02 | 0.03 | -0.08 – 0.05 | 0.625 |

| *Random Effects* | | |
|---|---|---|
| **$\sigma^2$** | 0.99 | 0.90 |
| $\tau_{00}$ | $0.60_{\text{SchoolID}}$ | $0.29_{\text{SchoolID}}$ |
| **ICC** | $0.06_{\text{District}}$ | $0.03_{\text{District}}$ |
| | 0.40 | 0.26 |
| **N** | $414_{\text{SchoolID}}$ | $413_{\text{SchoolID}}$ |
| | $57_{\text{Municipality}}$ | $56_{\text{Municipality}}$ |
| **Observations** | 32976 | 32946 |
| Marginal $R^2$ / Conditional $R^2$ | 0.000 / 0.399 | 0.254 / 0.449 |

## DISCUSSION

The examination of the two large-scale assessments suggests the existence of gender, community, and regional disparities in Albania. At the individual level, the PISA 2018 assessment shows that such disparities are also attributed to the socioeconomic status of the students. For the State Matura Exam, no data was available to study the effects of student characteristics, but similar gender disparities were observed. In the PISA 2018 dataset, gender and socioeconomic disparities are among the highest for Albania across all domains (e.g., Albania ranks 9$^{th}$ among 84 countries in reading regarding gender disparities in favor of girls; and 24$^{th}$ regarding socioeconomic disparities in mathematics, see: OECD 2019a, PISA 2018 parity index, Table 1.B1.50).

At the school level, our statistical model with the PISA 2018 dataset did not indicate that inequalities are due to the shortage of school resources. The model also shows that differences in public spending across districts are not associated with student achievement in the State Matura Exam. Although the findings are consistent with studies elsewhere that show inconclusive results (Della Sala, Knoeppel, & Marion, 2017; Hanushek, 1989; Hanushek & Woessmann, 2014), public spending on education has remained at three percent of Albania's GDP, which policymakers believe to be insufficient (Ministry of Education, 2014a). The expenditures per student at the secondary level are judged by international experts to be inadequate and comparably less than in other countries (UNESCO, 2017).

The results of the analysis of PISA and SME datasets show that inequalities across school types and community types are persistent in Albanian education. Public schools perform rather poorly. Given the significance of socioeconomic background from the PISA study, this could be attributed to clustering as students of higher socioeconomic status are enrolled mainly in private schools. The effects

of the socioeconomic status at the individual level are positively associated with the test achievement in PISA 2018. There is no data on SME to observe the same, but the literature suggests that students from wealthier families have the financial means to make use of private tutoring. In a survey conducted by IDRA (2012), the majority of students (56 percent) were benefiting from some type of private tutoring. As stated by UNESCO concerning the Albanian school system, private tutoring "has the potential to considerably distort the fundamental principles of equality in schools" and may cause the marginalization of a group of students (2017, p. 43). On the other hand, the effect of private instruction on student achievement remains empirically inconclusive (Ireson, 2004).

## CONCLUSION

Monitoring the quality of output and outcome in education has not been a policy priority in the Albanian education governance system. Although the government is committed to UNESCO's goals of Education for All (see the most recent strategy for the years 2014–2020, Ministry of Education, 2014a). Accountability and efficiency reports are focused mainly on input and access as input indicators. Educational indicators to evaluate schools' performance, e.g., school report cards, have been used only recently, while the evidenced disparities are yet to be problematized. As evidenced in other countries (see: UNESCO, 2018, pp. 11–12), education data in Albania are still incomplete.

Given the paucity of the data at the student level, it is suggested that the State Matura Examination may be used to produce output indicators to monitor the performance of schools and the education system (see: Fitz-Gibbon & Koch, 2000) and the achievement of UNESCO's sustainable development goals. As recommended by other international experts, Albania has to review the national examination system to improve its support mechanisms of evidence-based policymaking (Maghnouj et al., 2020). The monitoring system could be enhanced by surveying students on context variables that will inform the policymaking process and contain information on the quality of the system. Furthermore, SME's outcomes could be tied to a funding system based on the input-output performance indicators to offer policymakers a reliable monitoring instrument (see the discussion in: Bosker & Guldemond, 1991).

Comparative examination of the Albanian-speaking students abroad is another source of information. For example, students with migration backgrounds who have at least an Albanian parent receive similar or even lower results in PISA or

national exams in comparison to test-takers in Albania. In PISA 2018, most Albanian students scored below the second level in reading: respectively 52 percent in Albania, 78.5 percent in Kosovo, 66.5 percent in Montenegro, 63 percent in North Macedonia, and 72.3 percent in Switzerland (only in Greece there is 48 percent). Students in Albania performed higher than their country fellows in Greece, by almost 30 points in mathematics and 12 points in science. However, PISA 2018 average scores for the total sample were higher in Greece and Switzerland than in Albania, and the gender disparities in favor of girls were much lower (e.g., see the gender parity index of PISA 2018 in reading: Albania = 1.35, Greece = 1.22, Switzerland = 1.12, see: OECD 2019a, PISA 2018, Table 1.B1.50). Furthermore, another study conducted in the United Kingdom for students of different ethnic groups in Key Stage 2 and 4 with the dataset from the National Pupil Database (achievements of 2013), after adjusting for socioeconomic status, reported that Albanian speakers performed poorer than English native speakers (Strand, Malmberg, & Hall, 2015). While in Italy, in a survey conducted by the Istituto Nazionale di Statistica (ISTAT, 2018), about 30 percent of high school Albanian speaking students had been retained at least a grade (about 20 percent in the sample of 42,239 foreign students, one Albanian speaking parent, n = 2448). These results are counter-intuitive since one would expect higher scores even for those students with a migration background due to economic advantages and a better school environment. Whether these differences could be the effect of test-taking motivation (Silm et al., 2021) or socioeconomic factors such as parent's education and language spoken at home (Martins & Veiga, 2010), they provide another useful source of information and merit further investigation.

## References

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus? *Economics of Education Review, 21*(3), 245–262. DOI: 10.1016/S0272-7757(01)00006-1.

BalkanWeb (2015, June 1). Testi i Gjuhës dhe Letërsisë, maturantët protestë para MAS: Shumë pyetje s'kishin lidhje me tekstin [Language and Literature Test, Students Protest in Front of the Ministry of Education: Many Response Items Were Not Related with the Text]. *Gazeta Shqiptare*. Retrieved from: https://www.balkanweb.com/site/testi-i-gjuhes-dhe-letersise-maturantet-proteste-para-mas-shume-pyetje-skishin-lidhje-me-tekstin/.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: 10.18637/jss.v067.i01.

Borgonovi, F., Ferrara, A., & Maghnouj, S. (2018). *The Gender Gap in Educational Outcomes in Norway.* OECD Education Working Papers, 183. https://doi.org/10.1787/f8ef1489-en.

Bosker, R., & Guldemond, H. (1991). Interdependency of Performance Indicators: An Empirical Study in a Categorical School System. In: S.W. Raudenbush, & J. Douglas Willms

(Eds.), *Schools, Classrooms, and Pupils: International Studies of Schooling from a Multilevel Perspective* (pp. 101–113). San Diego, CA: Academic Press.

Chapman, C., Muijs, D., Reynolds, D., Sammons, P., & Teddlie, C. (Eds.) (2016). *The Routledge International Handbook of Educational Effectiveness and Improvement: Research, Policy, and Practice*. New York: Routledge.

Contini, D., Di Tommaso, M.L., & Mendolia, S. (2017). The Gender Gap in Mathematics Achievement: Evidence from Italian Data. *Economics of Education Review, 58*, 32–42. DOI: 10.1016/j.econedurev.2017.03.001.

Della Sala, M.R., Knoeppel, R.C., & Marion, R. (2017). Modeling the Effects of Educational Resources on Student Achievement: Implications for Resource Allocation Policies. *Education and Urban Society, 49*(2), 180–202. DOI: 10.1177/0013124516643757.

Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). School Quality and the Gender Gap in Educational Achievement. *American Economic Review, 106*(5), 289–295. DOI: 10.1257/aer.p20161074.

Fitz-Gibbon, C., & Koch, S. (2000). School Effectiveness and Education Indicators. In: C. Teddlie, & D. Reynolds (Eds.), *The International Handbook of School Effectiveness Research* (pp. 257–282). London: Falmer Press.

Fowler, W.J., & Walberg, H.J. (1991). School Size, Characteristics, and Outcomes. *Educational Evaluation and Policy Analysis*, 13(2), 189–202. DOI: 10.3102/01623737013002189.

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Goldstein, H., & Thomas, S. (1996). Using Examination Results as Indicators of School and College Performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 159*(1), 149–163. DOI: 10.2307/2983475.

Gortazar, L., Kutner, D., & Inoue, K. (2014). *Education Quality and Opportunities for Skills Development in Albania: An Analysis of PISA 2000–2012 Results*. Washington, D.C.: World Bank Group. Retrieved from: http://documents.worldbank.org/curated/en/2012/01/20214469/education-quality-opportunities-skills-development-albania-analysis-pisa-2000-2012-results.

Hanushek, E.A. (1989). The Impact of Differential Expenditures on School Performance. *Educational Researcher, 18*(4), 45–62. DOI: 10.2307/1176650.

Hanushek, E.A., & Woessmann, L. (2014). Institutional Structures of the Education System and Student Achievement: A Review of Cross-Country Economic Research. In: R. Strietholt, W. Bos, J.E. Gustafsson, & M. Rosen (Eds.), *Educational Policy Evaluation through International Comparative Assessments* (pp. 145–175). Munster: Waxman.

Hox, J.J., Moerbeek, M., & Van de Schoot, R. (2010). *Multilevel Analysis: Techniques and Applications* (2nd Ed.). New York: Routledge.

IDRA (2012). *Matura Shtetërore: Një Vlerësim i Cilësisë së Sistemit* [State Matura: A Quality Assessment of the System]. Tirana: Viti i Botimit. Retrieved from: https://www.idrainstitute.org/files/reports/State%20Matura/Matura%20Shteterore%202012.pdf.

IEA (2019). IEA International Database (IDB) Analyzer (Version 4.0.35) [Software]. Hamburg, Germany: IEA – International Association for the Evaluation of Education Achievement.

INSTAT (2018). *Arsimi* [Education]. Retrieved from: http://databaza.instat.gov.al/pxweb/en/.

Ireson, J. (2004). Private Tutoring: How Prevalent and Effective Is It? *London Review of Education*, *2*(2), 109–122. DOI: 10.1080/1474846042000229458.

ISTAT (2018, November 29). *Integrazione delle seconde generazioni: Microdati ad uso pubblico*. Retrieved from: https://www.istat.it/it/archivio/210741.

Jacobs, J.E., & Bleeker, M.M. (2004). Girls' and Boys' Developing Interests in Math and Science: Do Parents Matter? *New Directions for Child and Adolescent Development*, *106*, 5–21. DOI: 10.1002/cd.113.

Laukaityte, I., & Wiberg, M. (2017). Using Plausible Values in Secondary Analysis in Large-Scale Assessments. *Communications in Statistics – Theory and Methods*, *46*(22), 11341–11357. DOI: 10.1080/03610926.2016.1267764.

Lecki, G. (2010). LEMMA (Learning Environment for Multilevel Methods and Applications). Module 5: Introduction to Multilevel Modelling R Practical. Retrieved from: https://www.cmm.bris.ac.uk/lemma/.

Liu, J., Peng, P., & Luo, L. (2020). The Relation between Family Socioeconomic Status and Academic Achievement in China: A Meta-Analysis. *Educational Psychology Review*, 32(1), 49–76. DOI: 10.1007/s10648-019-09494-0.

Lockheed, M.E., Wagemaker, H. (2013). International Large-Scale Assessments: Thermometers, Whips or Useful Policy Tools? *Research in Comparative and International Education*, 8(3), 296–306. DOI:10.2304/rcie.2013.8.3.296

Lubienski, C.A., & Lubienski, S.T. (2013). *The Public School Advantage: Why Public Schools Outperform Private Schools*. Chicago: University of Chicago Press.

Lucas, G., Daniel, K., & Keiko, I. (2014). *Education quality and opportunities for skills development in Albania: an analysis of Pisa 2000–2012 results*. Retrieved from: http://documents.worldbank.org/curated/en/2012/01/20214469/education-quality-opportunities-skills-development-albania-analysis-pisa-2000-2012-results

Lüdecke, D. (2018a). *sjPlot: Data Visualization for Statistics in Social Science*. Retrieved from: https://cran.r-project.org/package=sjPlot.

Lüdecke, D. (2018b). *sjstats: Statistical Functions for Regression Models (Version 0.17.1)*. Retrieved from: https://CRAN.R-project.org/package=sjstats.

Lunn, D.J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian Modelling Framework: Concepts, Structure, and Extensibility. *Statistics and Computing*, *10*(4), 325–337. DOI: 10.1023/A:1008929526011.

Maghnouj, S., Fordham, E., Guthrie, C., Henderson, K., & Trujillo, D. (2020, May 7). *OECD Reviews of Evaluation and Assessment in Education: Albania*. Paris: OECD.

Martins, L., & Veiga, P. (2010). Do Inequalities in Parents' Education Play an Important Role in PISA Students' Mathematics Achievement Test Score Disparities? *Economics of Education Review*, *29*(6), 1016–1033. DOI: 10.1016/j.econedurev.2010.05.001.

Meyer, R.H. (1993). Can Schools Be Held Accountable for Good Performance? A Critique of Common Educational Performance Indicators. In: E.P. Hoffman (Ed.), *Essays on the Economics of Education* (pp. 75–110). Kalamazoo: WE Upjohn Institute for Employment Research. DOI: 10.17848/9780880995542.ch4.

Ministry of Education and Science (2011). *School Inventory and Education Survey: Final Results*. (TA-ALB-10). Government of Albania.

Ministry of Education (2014a). *Dokumenti i strategjisë së zhvillimit të arsimit parauniversitar 2014–2020* [Document Strategy for the Development of Pre-University Education

2014–2020]. Tirana: Ministria e Arsimit dhe Sportit. Retrieved from: http://arsimi.gov.al/wp-content/uploads/2018/08/Strategji_APU_dokumenti_perfundimtar_24_03_2015.pdf.

Ministry of Education (2014b, February 14). *Karta e Performances se Shkolles* [School Performance Card].

National Agency of Examinations (2014). *Vlerësim Kombëtar (Për nxënësit e klasës III dhe V)* [National Assessment: For Students at the 3rd and 5th Grade]. Retrieved from: http://qsha.gov.al/infos/publikime_raporte/raport_komb_kl3-5.pdf.

Neely, S.R., & Diebold, J. (2016). Public Expenditures and the Production of Education. *Education Policy Analysis Archives*, *24*(88), 1–22. DOI: 10.14507/epaa.24.2441.

OECD (2010). *PISA 2009 Results: What Makes a School Successful?: Resources, Policies and Practices*. Paris: OECD. Retrieved from: https://www.oecd.org/pisa/pisaproducts/48852721.pdf.

OECD (2016, December 6). *PISA 2015 Results*. Paris: OECD.

OECD (2019a). *PISA 2018 Assessment and Analytical Framework*. Paris: OECD. DOI: 10.1787/b25efab8-en.

OECD (2019b). *Program for International Student Assessment (PISA): Results from PISA 2018: Albania*. Paris: OECD. Retrieved from: https://www.oecd.org/pisa/publications/PISA2018_CN_ALB.pdf.

OSFA (2013). *Matura Shtetërore ne Shqipëri*, 2012. Tirana: Fondacioni Shoqëria e Hapur për Shqipërinë.

Penner, A.M., & Paret, M. (2008). Gender Differences in Mathematics Achievement: Exploring the Early Grades and the Extremes. *Social Science Research*, *37*(1), 239–253. DOI: 10.1016/j.ssresearch.2007.06.012.

Prenzel, M. (2012). Introduction: Research on PISA, with PISA, and for PISA. In: M. Prenzel, M. Kobarg, K. Schöps, & S. Rönnebeck (Eds.), *Research on PISA: Research Outcomes of the PISA Research Conference 2009* (pp. xiii-xxii). Heidelberg: Springer Science & Business Media.

QSHA (2014). *Vlerësim Kombëtar (Për nxënësit e klasës 3 dhe 5)* [National Assessment: Students of the 3rd and 5th Grade]. Retrieved from: http://qsha.gov.al/infos/publikime_raporte/raport_komb_kl3-5.pdf.

QSHA (2017a). *Matura Shteterore 2017: Raport Publik mbi Arritjet e Nxenesve* [State Matura 2017: Public Report on Students' Achievements]. Retrieved from: http://qsha.gov.al/infos/publikime_raporte/raporti_msh_2017.pdf.

QSHA (2017b). *Provimet Kombetare te Arsimit Baze 2017.* Retrieved from: http://qsha.gov.al/infos/publikime_raporte/raporti_pkab_2017.pdf.

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raudenbush, S.W., Bryk, A., & Congdon, R. (2013). HLM 7.01 for Windows [Hierarchical Linear and Nonlinear Modeling Software]. Skokie, IL: Scientific Software International.

Raudenbush, S.W., & Willms, J.D. (Eds.) (2014). *Schools, Classrooms, and Pupils: International Studies of Schooling from a Multilevel Perspective*. San Diego, CA: Academic Press.

Rizzuto, R., & Wachtel, P. (1980). Further Evidence on the Returns to School Quality. *Journal of Human Resources*, *15*(2), 240–245. DOI: 10.2307/145333.

Sakellariou, C. (2017). Private or Public School Advantage? Evidence from 40 Countries Using PISA 2012-Mathematics. *Applied Economics*, *49*(29), 2875–2892. DOI: 10.1080/00036846.2016.1248361.

Sander, W. (1993). Expenditures and Student Achievement in Illinois: New Evidence. *Journal of Public Economics*, *52*(3), 403–416. DOI: 10.1016/0047-2727(93)90043-S.

Scheerens, J. (2016). *Educational Effectiveness and Ineffectiveness: A Critical Review of the Knowledge Base*. Dordrecht: Springer.

Scheerens, J., Hendriks, M., & Luyten, H. (2014). School Size Effects: Review and Conceptual Analysis. In: H. Luyten, M. Hendriks, & J. Scheerens (Eds.), *School Size Effects Revisited* (pp. 7–40). Dordrecht: Springer.

Schmidt-Neke, M. (2007). Albania. In: W. Hörner, H. Döbert, B. von Kopp, & W. Mitter (Eds.), *The Education Systems of Europe* (pp. 11–32). Dordrecht: Springer.

Serra, D., Barr, A., & Packard, T.G. (2011). *Education Outcomes, School Governance and Parents' Demand for Accountability: Evidence from Albania*. Policy Research Working Paper, 5643. Washington: World Bank. Retrieved from: https://openknowledge.worldbank.org/bitstream/handle/10986/19879/WPS5643.pdf?sequence=1&isAllowed=y.

Silm, G., Must, O., Täht, K., & Pedaste, M. (2021). Does Test-Taking Motivation Predict Test Results in a High-Stakes Testing Context? *Educational Research and Evaluation*, 1–27. DOI: 10.1080/13803611.2021.1949355.

Sirin, S.R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, *75*(3), 417–453. DOI: 10.3102/00346543075003417.

Somers, M.A., McEwan, P.J., & Willms, J.D. (2004). How Effective Are Private Schools in Latin America? *Comparative Education Review*, *48*(1), 48–69. DOI: 10.1086/379841.

Strand, S., Malmberg, L., & Hall, J. (2015, January 29). *English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database*. London, Oxford: Education Endowment Foundation, University of Oxford. Retrieved from: https://ore.exeter.ac.uk/repository/bitstream/handle/10871/23323/EAL_and_educational_achievement2.pdf?sequence=1.

Teddlie, C., Stringfield, S., & Reynolds, D. (2000). Context Issues within School Effectiveness Research. In: C. Teddlie, & D. Reynolds (Eds.), *The International Handbook of School Effectiveness Research* (pp. 160–185). London: Falmer Press.

Thomson, S. (2018). Achievement at School and Socioeconomic Background: An Educational Perspective. *npj Science of Learning*, *3*(1), 5. DOI: 10.1038/s41539-018-0022-0.

UNESCO (2017). *Albania: Education Policy Review: Issues and Recommendations*. Retrieved from: http://unesdoc.unesco.org/images/0025/002592/259245e.pdf.

UNESCO (2018). *Handbook on Measuring Equity in Education*. Montreal, Canada: UNESCO Institute for Statistics. Retrieved from: http://uis.unesco.org/sites/default/files/documents/handbook-measuring-equity-education-2018-en.pdf.

World Bank (2005, March). *Albania. Poverty and Education in Albania: Who Benefits from Public Spending*. Retrieved from: https://documents1.worldbank.org/curated/en/582031468004830192/pdf/319830AL.pdf.