DETERMINANTS OF STUDENTS' PROFICIENCY: SCHOOL FEATURES AND PERSONAL CHARACTERISTICS VS SES IN KOSOVO

Luigi Maria Solivetti

Abstract. This study tests whether the school context and students' personal characteristics improve male and female students' proficiency irrespective of the effects of families' socioeconomic inequalities. The object of the investigation is a country presenting relatively poor educational results: Kosovo.

The study used the PISA 2022 dataset and focused on the math scores of female and male Kosovo students. The statistical analyses employed balanced repeated replication regression models by students' gender.

The results show that SES is a significant but not unmitigated determinant of the students' proficiency. School features affect the students' scores, while students' personal features have an even larger impact on them. Girls' math scores are equivalent to boys' – which is uncommon – but are differently affected by the proficiency determinants.

These findings suggest that school features and personal traits change students' proficiency even when pupils, on average, come from a context negatively affecting their capabilities and present relatively low proficiency. Moreover, averagely lower socioeconomic conditions and the probable persistence of traditional gender roles are not synonymous with the gender gap in proficiency. Still, girls and boys are different in how they respond to proficiency factors, and therefore gender-specific policies are advisable to improve their proficiency.

1. Introduction

The Programme for International Student Assessment (PISA) is an OECD survey designed to assess worldwide education systems. PISA evaluates how well 15-year-old students can use their reading, mathematics, and science knowledge to tackle real-world problems. The 2022 survey (the most recent one) included 80 countries. For each student, PISA recorded more than one thousand variables measuring his/her proficiency, personal opinions and attitudes, and family socioeconomic characteristics. Moreover, PISA collected two other datasets containing hundreds of variables: one about the schools' features and the other about teachers' qualities, opinions, and attitudes.

Since the first PISA survey in 2000, numerous studies have tried to predict students' proficiency scores using micro and macro determinants. Scholars have particularly focused on three domains: the student's socioeconomic status (hereafter

SES), the school features and teachers' qualifications and attitudes, and the student's personal characteristics.

In the present study, we intend to test the impact of determinants belonging to the said three domains on the student's proficiency in a country whose students scored relatively low in the PISA 2022 survey: Kosovo. There has been a high interest in countries presenting poor proficiency results (Okpala, Okpala and Smith, 2010; Woldemichael, Semela and Tulu, 2023). Indeed, identifying the factors associated with lower educational results is crucial to reducing international inequalities in knowledge and development. Besides, comparing SES determinants of proficiency to other factors would be potentially more constructive when the country under investigation is a relatively poor one, such as Kosovo, which has the lowest per capita GDP in Europe (in 2024, 16.7K). Indeed, it would be essential to investigate whether, in less developed countries, non-material proficiency factors counterbalance the effects of an averagely unfavourable SES.

Moreover, it would be stimulating to check the impact of gender on proficiency in an adverse socioeconomic context. All over the world, girls perform, on average, better than boys in reading but worse in math and science, the branches of learning leading to better-paid and prestigious jobs. We expect a substantial gender gap in educational outcomes in less developed societies, owing to the persistence of traditional gender roles. There is, indeed, an inverse cross-country association between GDPs per capita in PPP and the gender gaps in math (Anghel, Rodríguez-Planas and Sanz-de-Galdeano, 2020). However, this association evaporates when we consider only non-OECD countries. PISA 2022 data show that none of the few countries presenting neatly higher math scores for girls belongs to the group of most advanced countries. In addition, regardless of the gender gap in proficiency, it would be worth investigating other potential dissimilarities revolving around gender, such as the different impacts that the same factors could have on girls' and boys' scores.

2. Review of the literature and present study's hypotheses

Current literature has shown great interest in the students' SES. Cross-country, a close correlation (r = 0.73) exists between average GDP per capita and students' average proficiency (World Bank, 2023). However, the correlation for the lowest GDP tertile countries is weaker than that for the highest tertile countries: r = 0.21 vs 0.32. The relationship between students' SES and proficiency is also positive at the micro level. For all the students in PISA 2022, we found a correlation of 0.48 between SES and math. The correlation is distinctly weaker for the students in the group of countries with lower socioeconomic conditions, although in these countries, the students' SES relative standard deviation is – as expected – higher. This suggests that, in underprivileged countries, the relationship between SES and proficiency is not only weaker but also less affected by inequalities in SES.

As a trend, a higher SES would entail more expenditures in goods and opportunities capable of increasing the children's human capital. This, in turn, would favour better educational results (Okpala, Okpala and Smith, 2010). In a higher SES house, there would be more traditional cultural goods, such as newspapers, magazines and books, and more new tools, such as Internet access, computers, and cell phones: tools that tend to increase students' performances (Schleicher, 2023: 34). High-SES students would also benefit from the opportunities created by their parents' expenditures on non-material goods: expenditures on their children's learning and skills, such as private tutoring, better schools etc. Besides, higher SES usually entails a higher level of parental involvement in their children's studies as well as attitudes that encourage children to pursue higher educational achievements. In turn, students from low-income families would be negatively affected not only by their inadequate learning environments at home but also by the pressure to contribute to household income (Post and Pong, 2009; Cooper and Miralay, 2022).

Ultimately, literature has ascribed the impact of the family's SES on the student's proficiency to various specific factors. Moreover, no consensus has emerged about how to measure the various elements underlying the SES acronym: social class, poverty, affluence, power, and culture (Pokropek, Borgonovi and Jakubowski, 2015). Numerous variables (from income to possessions, education, occupation and attitudes) and also various combinations of variables have been used for this purpose. For instance, PISA 2022 used a composite index (ESCS) derived from three indicators: parents' education, parents' occupational status and home possessions.

Secondly, the student's proficiency is associated with his/her experiences at school. Teacher-related factors – such as staff shortages, teacher absenteeism, teachers' qualifications, commitment, and attitudes – would impact the student's learning (Ammermüller, 2004). The same would occur with schools' characteristics, such as rural/urban, private/public, their availability of libraries and other material resources, and their strictness (Shukakidze, 2013). The substandard level of school infrastructures and teachers' working conditions and salaries seems to have affected Kosovari students' proficiency (Lutfiu and Hoxha, 2024). Moreover, the school's social environment would also affect the student's proficiency: the school's social context generates interactions that, in turn, shape the student's perception of self and sense of belonging. For instance, facts such as being bullied negatively influence students' proficiency (Schleicher, 2023).

In lower-income countries, school-related factors are expected to have a deeper impact on students' proficiency. In wealthier countries, students can benefit from educational materials and a good learning environment at home, as well as from private tutoring. In poorer countries, instead, school is the primary source of the student's learning opportunities (Shukakidze, 2013). Ultimately, the chances of educational success for students from an underprivileged background often depend

on a good teacher and a good school. However, one should consider that a school's quality is not independent of the student's SES. As we already remarked, higher-SES families tend to cherry-pick their children's schools. Therefore, an analysis of proficiency based on the school's and teachers' qualities would miss the indirect influence of the student's SES.

Thirdly, students' personal features cannot but influence their school performances. Non-cognitive attributes, namely skills and characteristics concerning the student's attitudinal, behavioural, emotional, and motivational dispositions, seem particularly momentous. Hattie (2009) found that "engagement and motivation", "self-concept", and "anxiety" were the most relevant ones. In turn, other scholars emphasised the relevance of "performance self-efficacy", e.g. the perception of one's performance capability, in predicting educational scores (Lee and Stankov, 2018). Other individual factors can influence proficiency. For instance, student labour can negatively impact educational scores (Post and Pong, 2009).

Given all of this, we advance the following hypotheses:

H1. Even within a country with an averagely low proficiency, there is substantial variance in the students' scores: the precondition for further investigations.

H2. In such a country, it is possible to identify a relationship between the family's SES and the student's proficiency, but other factors also correlate with proficiency.

H3. School-related factors and the student's personal characteristics affect proficiency, acting as forces partially independent of SES.

H4. SES, school-related factors, and the student's personal characteristics can affect the proficiencies of girls and boys differently.

3. Methods

The present study is based on the PISA 2022 data (OECD, 2023) and, in particular, the Kosovo students' sample. We preliminarily merged the Student Questionnaire with the School and the Teacher Questionnaires. Kosovari students' proficiency has been our response variable. We focused on the scores in mathematics, the field of knowledge particularly apt to pass through cross-country differences in culture and language. For each student, PISA provides not a single score but ten plausible values as well as weights. To conduct statistical analyses on PISA proficiency data, it is therefore necessary to use specific procedures, as shown by the notation (1) for our main regression models:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \, [fw(w_i), rw(w_{1,i}, \dots, w_{nth,i}), vce(brr_i), fay(0.5)] \quad (1)$$

where y = plausible values for student i; x = explanatory variable; ε = error; fw = final survey weight incorporating any student-level trimming; rw = replicated weights (N = 80) for the balanced repeated replication (BRR) procedure taking into account sampling and non-response errors; vce = variance estimator for the said

procedure; *fay* = Fay's adjustment that perturbs the replicated weights by a specific per cent. The final results are the mean of the regressions of the ten plausible values. To conduct our analysis, we used the *repest* and *pisatools* applications for Stata.

As for the explanatory variables, we had to make a strict selection, owing to the vast number of measures concerning students' familial and personal characteristics, schools' features, and teachers' qualifications and attitudes. Firstly, we excluded all the variables without enough valid observations. We set the minimum threshold at 3500 valid observations, considering that we had to split the dataset into two parts, one for female students and the other for male students, and that using multiple variables in the regression models would inevitably multiply the final number of row-wise missing observations. Then, we excluded the redundant variables. After that, we checked the summary statistics to ascertain central tendencies and variances. Finally, we used a backward stepwise estimation procedure to filter the variables.

Literature on the determinants of proficiency has usually focused on a specific domain (e.g. the student's SES, or schools' and teachers' characteristics). Alternatively, it has analysed the effect on proficiency jointly made by determinants belonging to various domains. In the present study, we intend to follow a different method. Firstly, we will measure the effect of the family's SES on the student's proficiency. Then, we will analyse the impact of factors belonging to the school domain and to the student's personal characteristics. However, we know that the quality of the school is not independent of the student's SES, and there are links between the students' features (e.g. time for studying) and their familial background. Therefore, to avoid underestimating the direct and indirect effects of SES on proficiency, we will analyse the impact of the school and the student's personal characteristics on the proficiency variance unexplained by familial background. By doing this, we expect it would be possible to measure the impacts of the school and student's characteristics *at parity of family and socioeconomic determinants*.

4. Results

In the PISA 2022 tests, Kosovo's students obtained rather analogous scores in mathematics, reading, and science. Concurrently, those scores were markedly lower than the average scores registered for all the world countries, the European countries and even the Balkan countries (Figure 1).

Figure 1 also shows that the Kosovari girls' scores in reading were neatly higher than boys', which is in tune with the trend in the world's countries. Kosovo's girls also scored better than boys in science, while, on average, in the world, girls' scores are slightly lower than boys'. The gap in math between Kosovo's boys and girls, in turn, was almost non-existent, in contrast with what occurs elsewhere. Figure 2 shows two other facts. The math ten plausible values distributions follow the same pattern, not too dissimilar from a normal distribution. Therefore, Kosovari students'

math scores were widely different: their coefficient of variation (σ/μ) is 0.18, not much lower than all Europe's, which is 0.20. In conclusion, nothing hinders a statistical analysis of these distributions to identify the determinants of the variance in the students' scores. Concurrently, the girls' and boys' distributions of scores are not the same. While kurtosis is similar (3.1 vs 3.3), boys' scores show a greater positive skewness (0.52 vs 0.29). This gap is also found in other countries where, however, on average, it is smaller.

Figure 1 – *Kosovari boys' and girls' average proficiency scores compared to some world regions*'.

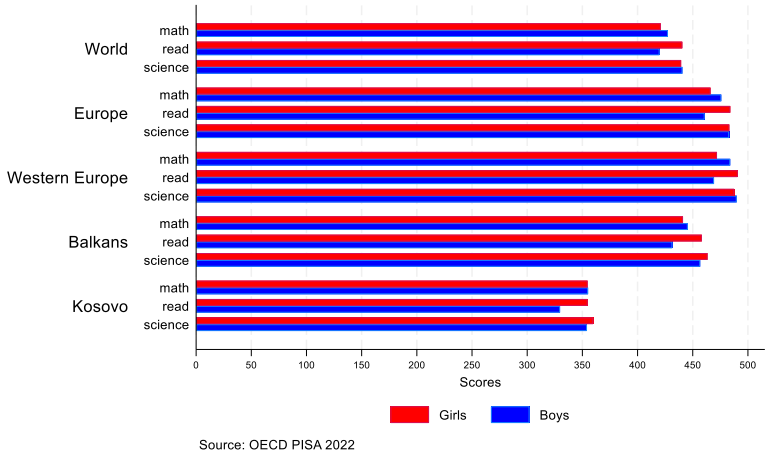
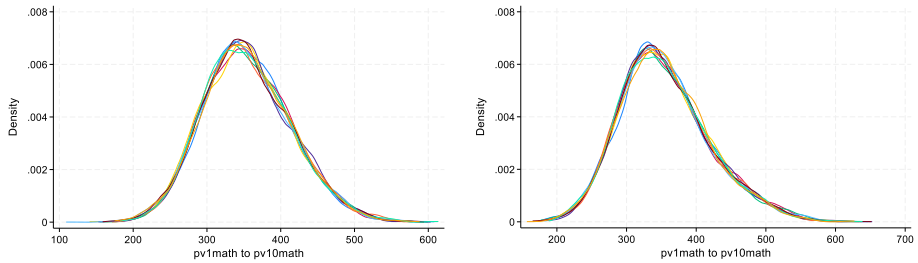


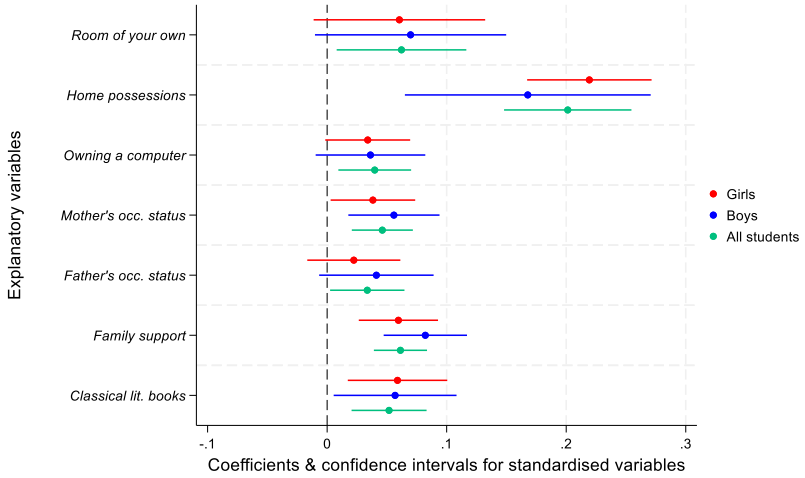
Figure 2 – *Kernel density estimates of Kosovari students' math plausible values.*



Girls Boys

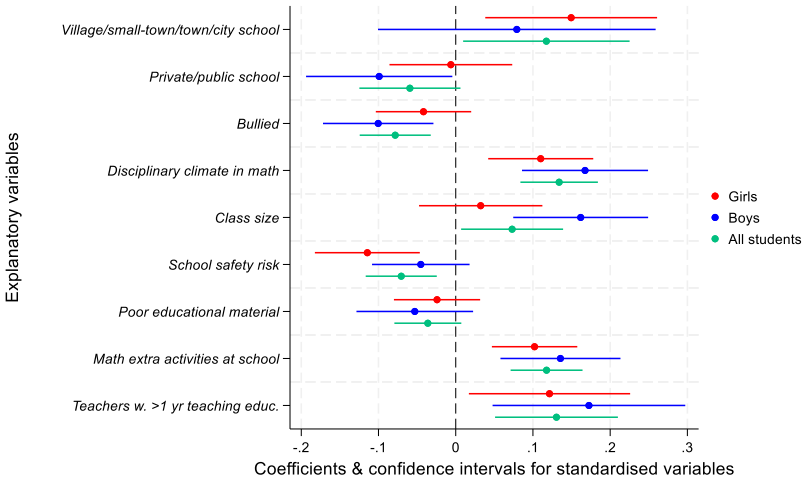
Note: kernel = Epanechnikov; weighted values

Figure 3 – *Balanced repeated replication regression by gender of Kosovari students' math scores on their families' socioeconomic determinants.*



Note: Girls N = 1706, R^2 = 0.14; Boys N = 1389, R^2 = 0.12; All students N = 3095, R^2 = 0.12

Figure 4 – *Balanced repeated replication regression by gender of the residuals of Kosovari students' math scores (from Figure 3 regression) on school determinants.*



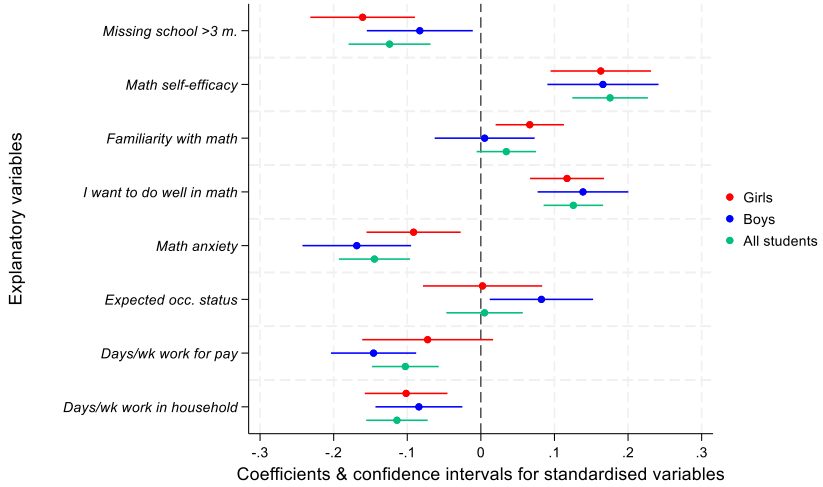
Note: Girls N = 1141, R^2 = 0.08; Boys N = 829, R^2 = 0.13; All students N = 1970, R^2 = 0.09

Figure 3 shows the results of the analysis of the socioeconomic determinants of Kosovari students' proficiency. All the seven explanatory variables directly correlate

with the math scores and are significant when considering girls and boys together. The father's occupational status is less impactful than the mother's. The impact of "home possessions" is particularly large in the case of girls. The models explain 12% to 14% of the proficiency variance.

Figure 4 shows the regressions by gender of the residuals of students' math scores (from the previous regression of Figure 3) on school determinants. The degree of urbanisation, the disciplinary climate in mathematics, the extra activities at school, and the teachers' education directly impact the students' proficiency. The same occurs with class size, which directly correlates with urbanisation. Instead, being bullied and the school safety risk are associated with lower scores. When we split the sample into two groups, girls and boys, it emerges that school urbanisation is significant only for girls, while being bullied significantly reduces only boys' proficiency. School safety risk significantly impacts only girls' proficiency, while being in a public school significantly reduces boys' proficiency. All in all, the school and teachers' characteristics explain 8% to 13% of the variance in proficiency left unaccounted by the SES determinants.

Figure 5 – *Balanced repeated replication regression by gender of the residuals of Kosovari students' math scores (from Figure 3 regression) on their personal determinants.*



Note: Girls N = 1307, R^2 = 0.15; Boys N = 1019, R^2 = 0.19; All students N = 2326, R^2 = 0.16

Figure 5 shows how the students' personal features impact the residuals of the regression of math scores on the family's socioeconomic determinants. Math self-efficacy and the students' determination to do well in math correlate with higher scores. For instance, every increase in the students' determination to do well in math

(4-level variable) averagely adds 14 points to their (residual) math scores. Instead, math anxiety negatively affects proficiency scores, and the same occurs with missing school for an extended span of time, working for pay and working in the household after school. We notice some gender differences: the expected occupational level is significant only for boys, while familiarity with math is significant only for girls. Besides, the anxiety impact on proficiency is smaller in the case of girls. In turn, working for pay only significantly impacts boys' proficiency. Apropos of this, we also found that, in Kosovo, male students work for pay significantly more than female students. Kosovari boys also work significantly more than boys in Western Europe. All in all, in Kosovo, personal characteristics explain 15% to 19% of the variance left unexplained by the socioeconomic determinants.

5. Discussion

Despite the comparatively low level of proficiency, the sample of Kosovo's students considered by PISA 2022 presents a well-differentiated distribution of their math scores. The distribution of the girls' scores shows a right tail less extended than the boys'. This feature is shared, on average, by all the countries but is more marked in Kosovo. This means that, among Kosovari girls, the incidence of top scorers is smaller.

As to the role of SES, our results confirm the findings of previous scientific literature as to its significant effect on school proficiency. Our regression model showed that SES measures explained 12% of the Kosovari students' variance in math scores. This figure is approximately halfway between the figures calculated by PISA for all the world countries, where SES accounts for ~5% to ~25% of the national variance in math scores. However, the variance explained by our model is two times the variance (~6%) explained by the PISA basic index of economic, social and cultural status (ESCS) when applied to Kosovo students.

This gap seems to derive from the determinants selected in the present study. While the importance of "home possessions" is evident – and this index is part of the wider ESCS index – "a room of your own" would measure the role of a student's *sacred space*, where one feels safe and comfortable. The specific relevance, among home possessions, of "a computer at home" derives from the opportunities for learning offered by such a device. Both the mother's and father's occupational status would measure the impact of class as something only partially overlapping with material well-being. However, it is noteworthy that the mother's occupational status emerged as more impactful than the father's: this confirms the relevance of the mother-child relationship for the younger generations' upbringing. In turn, classical literature books at home would gauge the family's cultural background as distinct from material well-being. Finally, "family support" would measure the social capital

on which the student can rely at home: something that can be independent of the family's economic status.

As for gender differences, the most relevant one is the larger impact of "home possessions" on the girls' proficiency. Because "home possessions" *per se* can hardly cause these gender differences, we believe that more copious "possessions" mirror non-observed familial socio-cultural factors favouring female education.

Despite these results, the variance explained by our SES model (12%) clearly leaves most of the Kosovari students' variation in proficiency unexplained.

The regression models based on the school's features explained a share of this residual variance, from 8% to 13%. The results highlight the relevance of both the teaching and the school context for the student's proficiency. As for teaching, the impact of variables such as longer teaching education, "disciplinary climate", and extra activities at school emerged. As for the school context, one can infer its importance from the impact of the school's urbanisation level, the fact of being bullied, and the school's safety risk. These models also detected noteworthy gender differences that helped better understand the school's role. Proficiency elasticity to safety risks is higher for girls, suggesting they are more exposed to such risks. Girls' proficiency is also more affected by the school's urbanisation level: this could be ascribed to the fact that the rural context would probably discourage female education. Lastly, being bullied is significant only for the boys' proficiency: this outcome draws attention to gender differences in interpersonal relations.

The regression models revolving around the student's personal characteristics explained a further share (from 15% to 19%) of the variance left unanswered by the SES variables. Math scores proved to be particularly elastic to variables concerning the student's self-confidence and commitment to study, such as math self-efficacy and determination to do well in math. Instead, working for pay and working in the household negatively impacted proficiency. Unsurprisingly, the same occurs with anxiety, inversely correlated with self-confidence as measured by math self-efficacy ($r = -0.34$). Anxiety about one's math proficiency had a larger impact on boys' scores. This conflicts with the fact that, on average, Kosovo's girls assert a higher level of anxiety than boys (0.34 vs 0.13): a difference common to the students of most countries. This contradiction suggests that either girls overrate their anxiety or this last is intrinsically different from boys' as to its effect on learning. In turn, girls' expected occupational status is higher than boys' but is less correlated with their determination to do well in math and has a non-significant impact on their proficiency, as if girls were less confident of their chances of obtaining a better status through better scores.

Ultimately, this study concerned students from the country with the lowest per capita income in Europe. This disadvantaged socioeconomic background seems to have affected the Kosovari students' capabilities, leading to results in proficiency

that, on average, are the lowest in Europe. However, we saw that, on an individual level, although the SES determinants impacted the students' proficiency, they left most of the proficiency variance unexplained.

By focusing on this variance unexplained by the SES determinants, this study was able to analyse the impact of school-related factors and students' personal characteristics on proficiency, bypassing the problem of the indirect effects of SES on proficiency. The regressions showed that schools' features reduced the amount of this unexplained variance, while the students' personal characteristics had an even larger impact on it. These findings attest that – despite a familial socioeconomic context that, on average, was unfavourable – school features and personal characteristics can change students' proficiency. The school emerged not only as the setting in which the family's SES is translated into proficiency – as suggested by current literature – but also as an institution capable of reworking this translation. This finding has important implications for government policies. The role played by the student's personal characteristics, in turn, seems crucial. In converting the students' SES into a personal realisation – as school proficiency is – these characteristics emerged from the present study as even more impactful than that SES. These findings represent an additional contribution to the current literature also because the impacts of school features and students' personal characteristics were calculated on the basis of proficiency values net of family's socioeconomic status.

As for the gender issue, the present results show that Kosovari girls' math scores were substantially equal to boys'. This outcome distinguishes Kosovo from most countries in the world, including those in Europe. Therefore, relatively lower socioeconomic conditions and the probable persistence of traditional gender roles are not synonymous with the gender gap in proficiency. Concurrently, Kosovari boys' proficiency is negatively affected by their greater involvement in work for pay. This fact distinguishes them from West European boys, who work significantly less and score in math neatly more than girls.

All in all, performances in educational achievements are much less closely associated with gender than with SES, personal characteristics and also school features. However, this does not mean that the gender issue is inconsequential. By analysing separately girls and boys, it has been possible to show that their proficiency scores differ in how they respond to factors commonly affecting educational outcomes. While boys' proficiency is particularly affected by being the victim of bullying, girls' is reactive to safety at school and urbanisation. Moreover, the girls' group is comparatively short of top scorers: a fact expected later to affect girls' chances of access to hard-to-obtain positions. These results suggest that any policy meant to improve educational outcomes should include gender-differentiated provisions in tune with the different impacts of specific conditions.

References

AMMERMUELLER A. 2004. PISA: *What makes the difference? Explaining the Gap in PISA Test Scores Between Finland and Germany*. Mannheim: Zentrum für Europäische Wirtschaftsforschung,

ANGHEL B., RODRÍGUEZ-PLANAS N., SANZ-DE-GALDEANO A. 2020. *Gender Equality and the Math Gender Gap*. SSRN Paper. Rochester, N.Y.: SSRN.

COOPER J.J., MIRALAY F. 2022. The Examination of Poverty Effect on Student Performance in Low-Income Countries, *Near East University Online Journal of Education*, Vol. 5, No. 2, pp. 37–46.

HATTIE J. 2009. *Visible Learning: A Synthesis of over 800 Meta-Analyses relating to Achievement*. London: Routledge.

LEE J., STANKOV L. 2018. Non-Cognitive Predictors of Academic Achievement, *Learning and Individual Differences*, Vol. 65, pp. 50–64.

LUTFIU B., LUTFIU HOXHA L. 2024. Socioeconomic Status of Teachers and Its Impact on Teaching Quality, *European Journal of Education and Pedagogy*, Vol. 5, No. 2, pp. 52–58.

OECD. 2023. *PISA 2022 Database*. Paris: OECD. https://www.oecd.org/en/data/datasets/pisa-2022-database.html

OKPALA C.O., OKPALA A.O., SMITH F.E. 2001. Parental Involvement, Instructional Expenditures, Family Socioeconomic Attributes, and Student Achievement, *The Journal of Educational Research*, Vol. 95, No. 2, pp. 110–115.

POKROPEK A., BORGONOVI F., JAKUBOWSKI M. 2015. Socio-Economic Disparities in Academic Achievement, *Learning and Individual Differences*, Vol. 42, pp. 10–18.

POST D., PONG S-L. 2009. Student Labour and Academic Proficiency in International Perspective, *Internat. Labour Review*, Vol. 148, No. 1-2, pp. 93-122.

SCHLEICHER A. 2023. *PISA 2022 Insights and Interpretations*. Paris: OECD.

SHUKAKIDZE B. 2013. Comparative Study: Impact of Family, School, and Students Factors on Students Achievements in Reading in Developed (Estonia) and Developing (Azerbaijan) Countries, *International Education Studies*, Vol. 6, No. 7, pp. 131–143.

WOLDEMICHAEL B., SEMELA T., TULU A. 2023. Mathematics Self-Efficacy, Self-Concept and Anxiety in Relation to Mathematics Achievement: The Case of Kafa Zone School Adolescents, Ethiopia, *F1000Research*, Vol. 12, 1494.

WORLD BANK, 2023. *Average Learning Outcomes* [dataset].

Luigi M. SOLIVETTI, Sapienza University of Rome, luigi.solivetti@uniroma1.it