Public Education Investment, Instructional Resources, and Student Achievement: Cross-National Evidence from PISA 2022 in the Context of Sustainable Development Goal 4

Reshma Shaik[1,*], Hanadi Osman Diab[2]

[1]Assistant professor, Sri Venkateswara University, India
[2]Instructor at Lebanese International University, Lebanon
Emails: reshmaarhaan2010@gmail.com; Hanadi.diab@liu.edu.lb

Abstract
All people must have access to educational opportunities which meet their needs through Sustainable Development Goal 4. The goal requires education systems to obtain sufficient funds and use their resources properly. The relationship between public education spending and student academic performance remains disputed because different countries achieve different results from their spending levels. The study employs PISA 2022 country-level scores which represent the first international assessment data published after COVID-19 to analyze public education expenditure as a GDP share together with pupil–teacher ratio and per-capita GDP in relation to student academic performance across three subjects. The study found that public education funding as percentage of GDP does not connect with PISA score results across 35 countries, showing no statistical link to tests ($r = -0.095$, $p = 0.586$). The pupil–teacher ratio serves as an effective predictor because it shows a strong negative relationship to student performance ($\hat{\beta} = -4.097$, $R^2 = 0.312$, $p < 0.001$). A three-variable regression model which combines expenditure share with pupil–teacher ratio and GDP per capita explains 59% of cross-country score variance ($R^2 = 0.592$). High-income economies dominate the upper achievement tier, but several upper-middle-income systems—notably Estonia and Poland—substantially outperform their GDP-predicted scores. The results show that organizations should focus their resources on developing teaching skills.

Keywords: PISA 2022; Education expenditure; Pupil–teacher ratio; Student achievement; SDG 4; Cross-national analysis; Education finance

1. Introduction

In December 2023, the OECD Programme for International student Assessment sent a clear message: out of the 37 OECD countries with PISA 2022, mean mathematics performance decreased by a record low 15 points, and reading decreased by 10 points, the largest consecutive decline in the history of this programme since its inception in 2000 [1]. The drop in reading is even older than COVID-19, but in mathematics, it seems like the first test to capture both decades of systemic trend and a crisis shock in a single test, thanks to the pandemic disruptions in mathematics [1]. The findings rekindle ancient debates regarding the motivations to student learning: where the money invested does not lead to the commensurate results, what is it at the system-level that actually leads to the differentiation in achievement across nations?

This query lies at the core of Sustainable Development Goal-4, which proposes to provide inclusive, equitable, and quality education to everyone by 2030 [2]. The development on SDG 4 indicators has lagged behind standards on almost all aspects. In 2024, global median public education spending was 4.2% of GDP, down from 4.3% in 2010, and 22% of countries had both achieved the international financing targets of 4% of GDP and 15% of government spending at the same time [3]. However, the correlation between spending rates and learning performance is not linear: OECD data consistently indicate that at the level of expenditure of less than a cumulative USD 75,000 per student during primary and secondary education, higher investment is associated with higher PISA scores, but beyond that level, the association is no longer effective [1]. Learning what constitutes success in the high-spending level, in which a majority of the OECD member countries will fall, is thus as essential as learning about the investment gap in low-income environments.

There are three items that this paper adds to this question. It offers a clean, repeatable cross-national study of the PISA 2022 results, based on just publicly available institutional data, and can be recreated with just a single Python script. It isolates the impact of the aggregate expenditure volume, instructional resource intensity (pupil teacher ratio) and economic development on achievement, revealing the three inputs to qualitatively different associations with student performance. Lastly, it puts the findings into the SDG4 accountability framework to make inferences about the way education ministries and development banks must make the priority of their investments.

The paper has the following organisation. Section 2 is a review of the empirical literature involved. The conceptual framework is presented in section 3. Section 4 contains data and methodology. Section 5 reports results. Findings are discussed in Section 6. Section 7 concludes.

2. Literature Review

2.1. Education Expenditure and Learning Outcomes

Whether the increased spending on education yields better student outcomes or not is a question that has kept the economists of education busy over several decades and the answer to this question cannot be summarized easily. Quasi-experimental school finance reform studies in national settings generally report small positive returns: an increase in sustained per-pupil spending by USD 1,000/year over 10 years is correlated with an improvement of test scores by 0.12 to 0.24 standard deviations in the United States [1, 4]. The association is weaker at the international level. Cross-national studies have always revealed that the share of aggregate spending in GDP is a poor predictor of PISA performance after controlling income level [1, 4]. The intuition is simple in that high-income countries are able to spend more and do more at the same time due to the more institutional and human capital environments that enable learning, regardless of the marginal dollar expended on school budgets.

In a recent study of 20 countries, over 20112023, a DEA and Malmquist productivity analysis revealed significant heterogeneity in the efficiency of converting public education budgets into measurable outcomes, with Japan, Morocco, and Turkey having sustained productivity gains over the study time period despite varying levels of expenditure in 20112023 [4] This finding points out that governance quality and allocation decisions mediate the input output relationship in education- a fact that makes it difficult to come up with simple spending performance comparisons like those presented in SDG monitoring systems.

2.2. Instructional Resources: Teacher Quality and Pupil–Teacher Ratio

If aggregate spending volume is an imperfect predictor, the composition of that spending matters considerably. Teacher compensation typically accounts for 60–80% of education budgets in OECD countries, and both teacher qualification levels and pupil–teacher ratios have been shown to have more direct associations with student performance than aggregate expenditure totals [5]. Multilevel analyses of PISA 2022 data for Singapore, South Korea, Finland, and Denmark find that teacher support—particularly availability during school closures—is one of the strongest student-level predictors of mathematics performance, with teacher availability associated with score advantages of 15 points even after controlling for socioeconomic background [6]. The cross-national evidence for pupil–teacher ratio as a system-level predictor is consistent: smaller secondary classes create conditions for more individualised instruction, and their positive association with achievement holds across diverse income settings [7].

2.3. Income, Socioeconomic Background, and the SDG 4 Equity Agenda

The best single predictor of PISA performance at the country level is GDP per capita, which is to a large extent a proxy of the wider socioeconomic context within which schooling is situated in its contextualization and interpretation [1]. In 2022, students with socio-economic advantages performed significantly higher on average in the 93 PISA points across OECD nations compared to their disadvantaged counterparts on average (about three years of learning). However, this average masks a significant range: in Macao (China), poor students in the system scored above the OECD average on average and even above average in the top-tier, indicating that high levels of system-wide equity can be obtained even in the absence of high GDP per capita. Equally, Estonia and Poland are long-standing outperformers on their income-predicted performances and their experience, including high teacher professionalism, late ability tracking, and high curriculum standards, is more and more proposed as transferable models to middle-income reform [1, 6].

The literature on SDG4 financing points out that bridging the global learning gap will need an increased investment in underfunded low-income environments and a more efficient distribution of the current resources in the middle- and high-income systems [3]. An annual financing gap estimated at USD 97 billion in 79 low- and lower-middle-income countries in 2023 was calculated to achieve SDG 4 targets; the largest financing shortfall was in sub-Saharan Africa [5]. To fill this gap with scaled-up investment, it needs to be evidenced what to invest in to get the greatest returns per dollar, which is the key empirical question of this paper.

2.4. Education, Sustainability, and the SDG 4 Nexus

Education systems and their intersection with larger sustainable development goals have been the focus of increasing empirical interest. Based on quantile regression of G20 countries post-2000 to 2022, [8] report that the educational investment by the government has a large positive impact on primary school enrolment at the low end of the enrolment distribution, whereas environmental and economic factors, such as access to renewable energy and financial inclusion, moderate the relationship between the educational investment and the enrolment This fact supports the argument of treating SDG 4 as a dependent and not a stand-alone goal in the 2030 Agenda, and encourages the focus of the present paper in multi-variable models, instead of single-input analyses.

3. Conceptual Framework and Hypotheses

Figure 1 presents the conceptual model organising this study. Public education finance constitutes the enabling resource base through which governments deploy instructional inputs. Three primary channels are hypothesised.

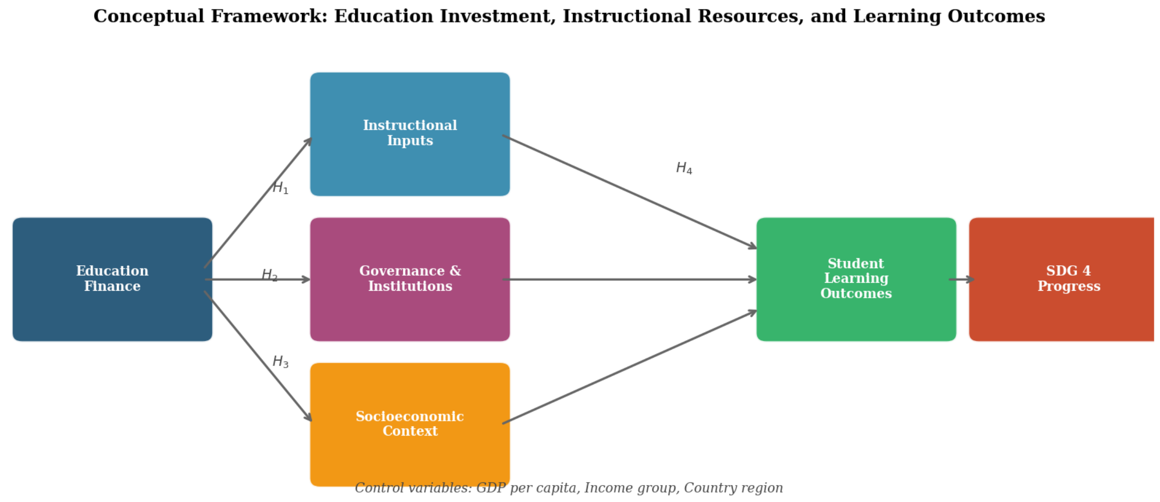

Figure 1: Conceptual framework linking public education finance to student learning outcomes via instructional inputs (H_1), institutional governance (H_2), and socioeconomic context (H_3). Together, these pathways are hypothesised to explain cross-national variation in PISA achievement (H_4). Control variables are listed at the base of the diagram.

The *instructional input channel* (H_1) links expenditure to deployment of teaching staff: higher spending, when efficiently allocated, should reduce class sizes and improve teacher qualification rates, with a direct route to student outcomes. The *governance and institutional channel* (H_2) captures the ways in which system-level decisions—curriculum design, assessment policy, teacher evaluation, and school autonomy arrangements—translate financial inputs into pedagogical practice. The *socioeconomic context channel* (H_3) acknowledges that learning outcomes reflect not only in-school resources but the broader material and cultural environment in which students learn.

Four testable hypotheses follow. H_1: countries with lower pupil–teacher ratios achieve higher mean PISA scores. H_2: public education expenditure as a share of GDP is positively associated with composite PISA scores, net of income effects. H_3: GDP per capita is positively associated with composite PISA scores. H_4: a model combining pupil–teacher ratio, education expenditure, and GDP per capita provides a statistically robust explanation of cross-country achievement variation.

4. Data and Methodology

4.1. Data Sources and Sample

The sample comprises 35 countries with complete data across all variables, drawn from OECD and non-OECD participants in PISA 2022. Countries were selected on the basis of data completeness for all four study variables; the final set spans high-income OECD members alongside several upper-middle and lower-middle income participants (Chile, Turkey, Mexico, Colombia, Brazil), providing cross-income variation essential for testing H_2 and H_3.

Four variables are employed. **PISA 2022 scores** in mathematics, reading, and science—and a composite mean of the three—are drawn from Table I.B1.1 of the OECD PISA 2022 Results, Volume I, released 5 December 2023 [1]. The PISA 2022 public dataset, including the full microdata file, is available at `https://www.oecd.org/en/data/datasets/pisa-2022-database.html`. **Public education expenditure as a share of GDP** is sourced from the UNESCO Institute for Statistics (UIS), series "Government expenditure on education as % of GDP", for the most recent available year in each country (predominantly 2019–2021) [9]. **Pupil–teacher ratio** in secondary education is from World Bank World Development Indicators series SE.SEC.STUDT.TCH.ZS [10]. **GDP per capita** (thousand USD, current prices, 2022) is from World Bank WDI series NY.GDP.PCAP.CD.

4.2. Descriptive Statistics

Table 1 reports summary statistics for all study variables. Mean composite PISA score across the 35 countries is 482.4 points (SD = 33.1), with a range from 397.3 (Brazil) to 559.7 (Singapore). Mean public education expenditure is 5.3% of GDP (SD = 1.4), ranging from 2.7% (Singapore) to 7.9% (Norway and Iceland). Pupil–teacher ratios range from 8.4 (Lithuania) to 28.6 (Mexico), with a mean of 13.7. GDP per capita spans USD 6,900 (Colombia) to USD 106,600 (Norway).

Table 1: Descriptive statistics, cross-sectional sample ($n = 35$)

Variable	Mean	SD	Min	P25	Median	Max
Mathematics score (pts)	476.5	38.5	379	470	482	575
Reading score (pts)	481.3	28.4	409	474	480	543
Science score (pts)	489.4	34.3	403	477	492	561
Composite PISA score (pts)	482.4	33.1	397	475	486	560
Education expenditure (% GDP)	5.3	1.4	2.7	4.3	5.5	7.9
Pupil–teacher ratio (secondary)	13.7	4.5	8.4	10.9	12.3	28.6
GDP per capita (k USD, 2022)	45.4	26.4	6.9	25.7	45.9	106.6

SD = standard deviation; P25 = 25th percentile.

Sources: OECD PISA 2022 Results Vol. I; UNESCO UIS; World Bank WDI.

4.3. Model Specification

Three bivariate OLS models and one multiple regression are estimated.

4.3.1. Models 1–3: Bivariate OLS

Each bivariate specification takes the form:

$$\text{PISA}_i = \alpha + \beta_k\, X_{k,i} + \varepsilon_i \tag{1}$$

where PISA_i is the composite score for country i; X_k is one of the three predictors (education expenditure, pupil–teacher ratio, GDP per capita); and $\varepsilon_i \sim (0, \sigma^2)$. The OLS estimator is:

$$\hat{\beta}_k = \frac{\sum_{i=1}^{n}(X_{k,i} - \bar{X}_k)(\text{PISA}_i - \overline{\text{PISA}})}{\sum_{i=1}^{n}(X_{k,i} - \bar{X}_k)^2} \tag{2}$$

Statistical significance is assessed via the t-statistic $t = \hat{\beta}_k / \hat{\sigma}_{\hat{\beta}_k}$ against a $t(n-2)$ distribution. Goodness of fit is evaluated by:

$$R^2 = 1 - \frac{\sum_i(\text{PISA}_i - \widehat{\text{PISA}}_i)^2}{\sum_i(\text{PISA}_i - \overline{\text{PISA}})^2} \tag{3}$$

4.3.2. Model 4: Multiple regression

The joint model is:

$$\text{PISA}_i = \alpha + \beta_1\,\text{EduExp}_i + \beta_2\,\text{PTR}_i + \beta_3\,\text{GDP}_i + \varepsilon_i \tag{4}$$

with coefficients recovered via the normal equations $\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y}$. Pairwise Pearson correlations are computed over the full sample via the standard formula:

$$r_{XY} = \frac{\sum_i(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i(X_i - \bar{X})^2 \cdot \sum_i(Y_i - \bar{Y})^2}} \tag{5}$$

5. Results

5.1. Country-Level Achievement and System Inputs

Figure 2 presents a four-panel overview of the dataset. Panel (a) ranks all 35 countries by composite PISA score. Singapore leads by a wide margin (559.7 points), followed by Japan (533.0) and South Korea (523.3). Estonia (515.7) and Canada (506.3) are the highest-placed European and North American economies respectively, with the United Kingdom, Australia, and Germany clustered between 485 and 492. At the lower end, Brazil (397.3) and Colombia (398.7) stand well below the OECD mathematics average of 472. A clear stratification between high-income and other economies is visible, though several notable deviations—Estonia and Poland above their income band, Turkey within reach of the European mid-range despite much lower GDP—suggest that system factors beyond income matter considerably.

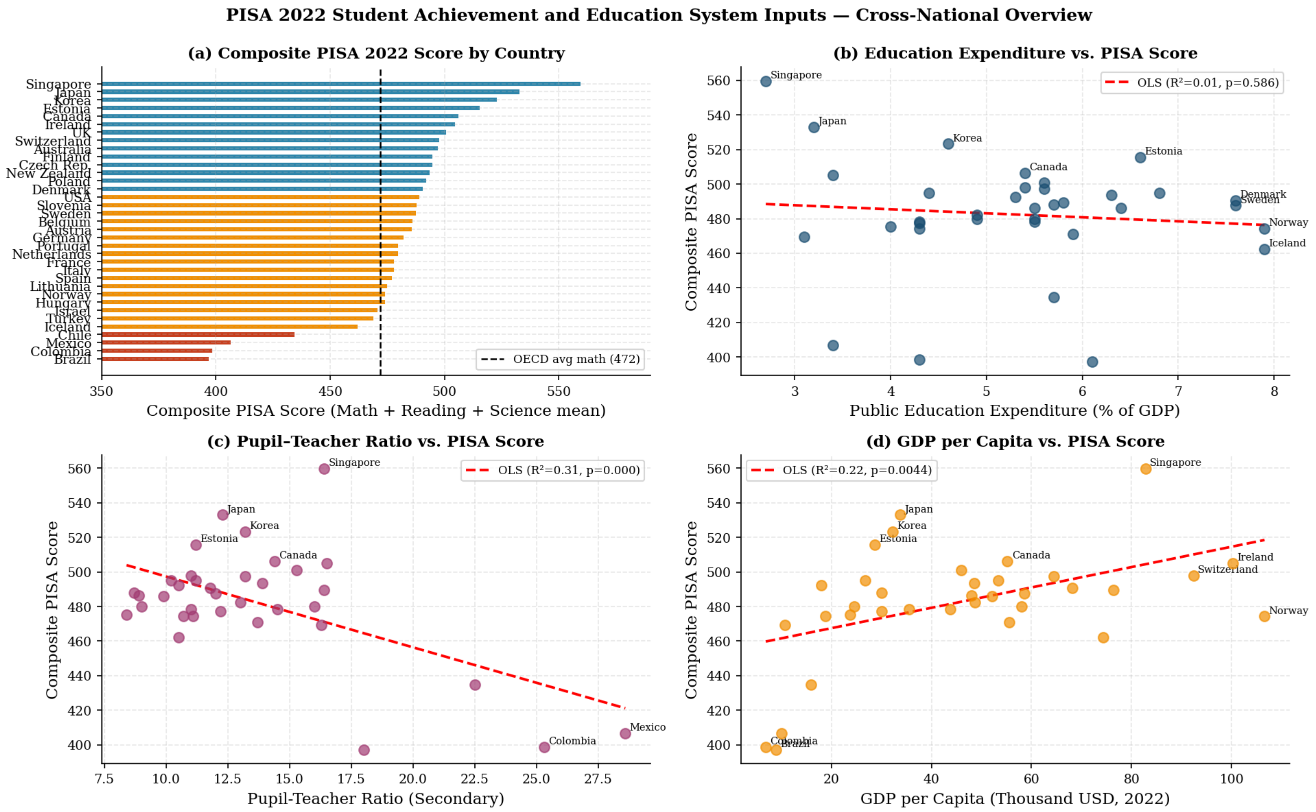

Figure 2: Four-panel overview of PISA 2022 outcomes and education system inputs. **(a)** Countries ranked by composite PISA score (colour bands: blue ≥ 490; orange 450–489; red < 450). **(b)** Education expenditure (% GDP) vs. composite score with OLS fit. **(c)** Pupil–teacher ratio vs. composite score with OLS fit. **(d)** GDP per capita vs. composite score with OLS fit.

Panel (b) reveals the absence of any meaningful positive association between education expenditure as a share of GDP and composite PISA performance. Countries with very high expenditure shares—Norway and Iceland at 7.9%, Denmark at 7.6%, Finland at 6.8%—achieve scores that are high but not disproportionate, while Singapore achieves the highest scores in the dataset on the lowest expenditure share (2.7% of GDP). The OLS slope is weakly negative ($\hat{\beta} = -2.32$), reflecting that some of the highest spenders in the sample are Scandinavian economies with strong but not exceptional PISA results.

Panel (c) shows a substantially stronger and clearly negative association between pupil–teacher ratio and PISA score. Systems with the smallest secondary class sizes—Belgium (8.9), Lithuania (8.4), Slovenia (8.7), and Austria (9.9)—achieve scores above the OECD average, while Mexico (28.6), Colombia (25.3), and Chile (22.5) combine high ratios with the lowest achievement levels in the sample. Panel (d) documents the expected positive income–achievement gradient, but also highlights meaningful departures: Estonia scores approximately 70 points above its GDP-predicted value, and Poland approximately 60 points above.

5.2. Score Distributions by Subject

Figure 3 presents violin and box plots of score distributions by subject. Mathematics exhibits the widest spread (SD $= 38.5$ points), consistent with the large within-sample variance introduced by Singapore's exceptional performance and by lower-income Latin American economies. Reading and science distributions are somewhat more compressed, with smaller outlier influence from the top. Across all three subjects, the OECD average falls at or above the sample median, reflecting that the 35-country sample includes several non-OECD economies with below-average scores.

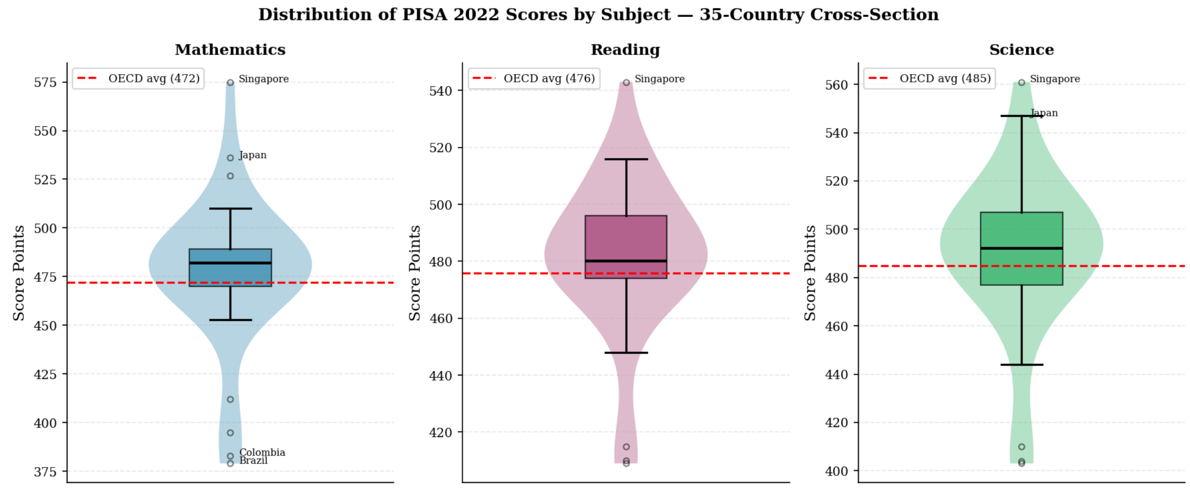

Figure 3: Violin and box plots of PISA 2022 score distributions for mathematics, reading, and science across 35 countries. Red dashed lines mark OECD averages (mathematics: 472; reading: 476; science: 485). Labelled points are notable high or low performers.

5.3. Model 1: Education Expenditure and Achievement

The bivariate OLS regression of composite PISA score on education expenditure as a share of GDP yields:

$$\widehat{\text{PISA}} = 494.8 - 2.32 \cdot \text{EduExp}, \quad R^2 = 0.009, \quad p = 0.586, \quad n = 35$$

The slope is negative and far from statistically significant ($t = -0.55$). The panel Pearson correlation is $r = -0.095$, confirming the absence of a meaningful linear association. Disaggregating by subject produces similarly negligible correlations: $r = -0.106$ for mathematics, $r = -0.082$ for reading, and $r = -0.089$ for science (all $p > 0.54$). This result directly supports the OECD's own finding that, within the high-expenditure tier, the level of investment is less predictive of outcomes than the efficiency and composition of how resources are deployed [1].

Figure 4 examines this relationship further. Panel (a) ranks countries by expenditure share, with high spenders ($\geq 5.5\%$ of GDP, $n = 18$) distinguished from lower spenders ($n = 17$). Panel (b) compares mean subject scores between these two groups and the OECD average. High spenders achieve a mean composite of 479.4 versus 485.6 for lower spenders—a difference of 6.2 points in the direction opposite to the investment hypothesis. While neither group differs significantly from the other or from the OECD average, the directional reversal points to a compositional effect: the high-spending group includes Scandinavian countries (strong performance) alongside Chile and Brazil (lower spending relative to their income level but in the bottom tier on PISA), while the low-spending group includes Singapore and Japan, whose fiscal conservatism in education coexists with top-tier outcomes.

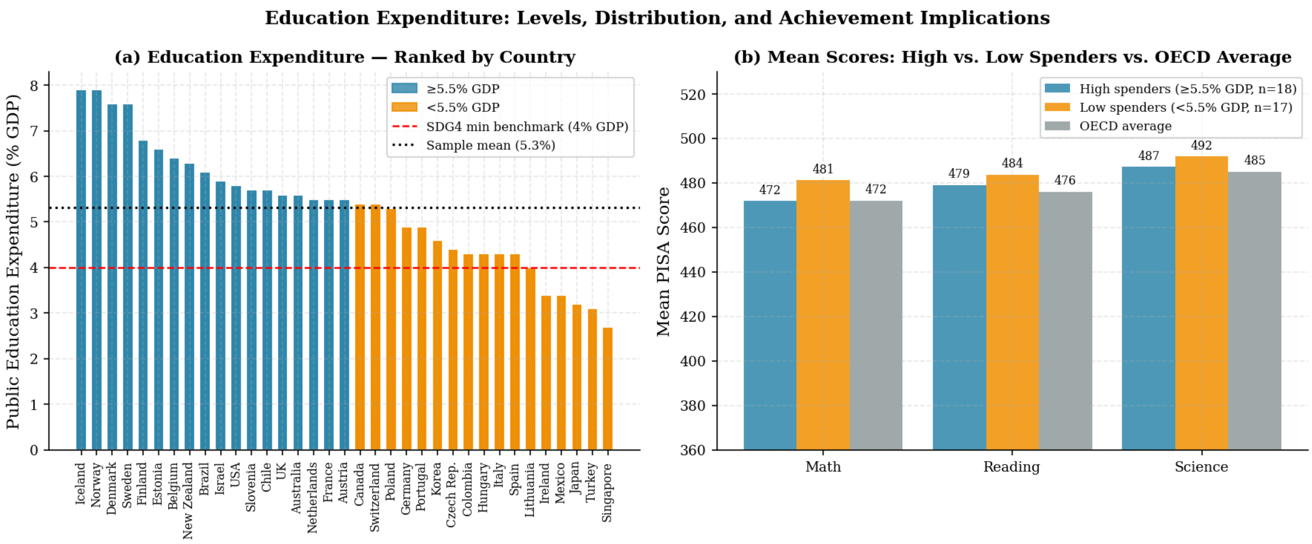

Figure 4: Education expenditure analysis. **(a)** Countries ranked by education expenditure share of GDP; blue bars indicate high spenders ($\geq 5.5\%$ GDP), orange bars indicate lower spenders. The red dashed line marks the SDG 4 minimum benchmark of 4% of GDP. **(b)** Mean subject scores for high spenders, lower spenders, and the OECD average (labelled values above bars).

DOI: https://doi.org/10.54216/IJAIET.040203

5.4. Model 2: Pupil–Teacher Ratio and Achievement

The bivariate regression of composite score on pupil–teacher ratio yields:

$$\widehat{\text{PISA}} = 538.4 - 4.097 \cdot \text{PTR}, \quad R^2 = 0.312, \quad p < 0.001, \quad n = 35$$

The slope is negative and highly significant ($t = -3.87$, $p = 0.0005$), indicating that each additional student per teacher is associated with a reduction of approximately 4.1 PISA points in the composite score. Applying this to the range between the sample's lowest and highest pupil–teacher ratios (8.4 to 28.6)—a difference of 20.2 students per teacher—implies an estimated score differential of approximately 83 points, comparable to nearly two full years of learning. Subject-level correlations are all significant: mathematics ($r = -0.579$, $p = 0.0003$), reading ($r = -0.509$, $p = 0.0018$), and science ($r = -0.546$, $p = 0.0007$).

5.5. Model 3: GDP per Capita and Achievement

The bivariate regression of composite score on GDP per capita yields:

$$\widehat{\text{PISA}} = 455.7 + 0.589 \cdot \text{GDP}, \quad R^2 = 0.221, \quad p = 0.004, \quad n = 35$$

The slope is positive and significant ($t = 3.06$), consistent with H_3. Figure 5(a) presents GDP-adjusted residuals—the gap between each country's actual and GDP-predicted composite score. Estonia (+56.4 points), Singapore (+55.3), and Poland (+47.8) are the strongest positive outliers, outperforming their income predictions substantially. Mexico (−44.9), Colombia (−37.5), and Brazil (−30.2) are the largest negative outliers, achieving below what their income level would predict.

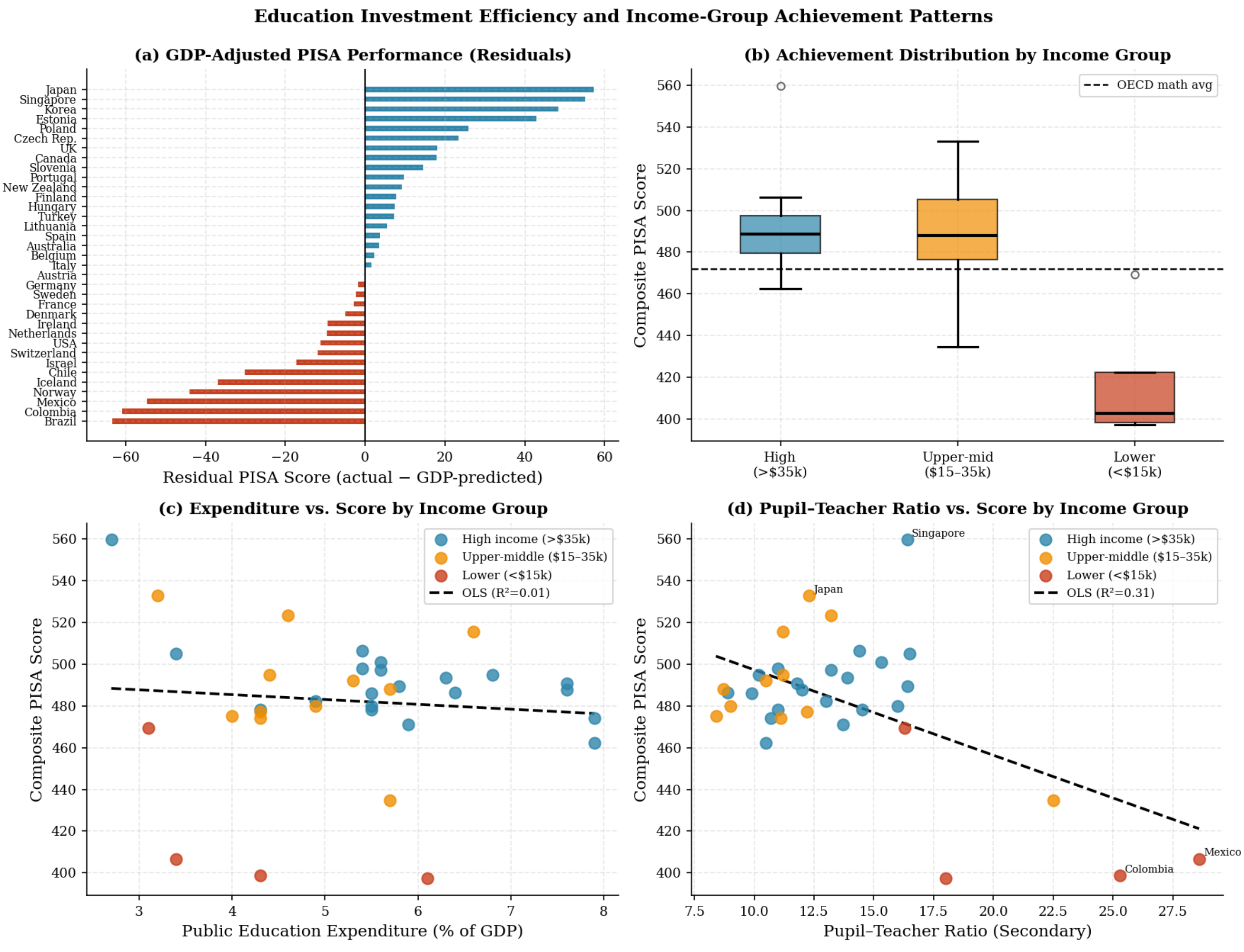

Figure 5: Income-adjusted achievement patterns. **(a)** Residuals from Model 3 (actual minus GDP-predicted composite score); positive values indicate overperformance relative to income. **(b)** Composite score distribution by income group (box plots). **(c)** Education expenditure vs. score, coloured by income group. **(d)** Pupil–teacher ratio vs. score, coloured by income group.

Panel (b) shows composite score distributions by income group. High-income economies (>USD 35k per capita) achieve a median composite of approximately 489 points, well above the OECD mathematics average. Upper-middle-income countries (USD 15–35k) show wider variance, with Estonia and Poland at the high end and the Czech Republic and Slovenia at the median. Lower-income economies (<USD 15k per capita) in the sample—Colombia, Brazil, Mexico, Turkey—cluster near the bottom.

Panels (c) and (d) replot the expenditure and PTR scatter plots with income groups colour-coded. A key observation in panel (c) is that within the high-income group, expenditure share and achievement are effectively uncorrelated: high-income countries span the full range from 2.7% (Singapore) to 7.9% (Norway) with scores concentrated between 480 and 560. Panel (d) confirms that the PTR–score relationship holds within each income group, suggesting that class size is an independent predictor of outcomes beyond what income alone explains.

5.6. Correlation Structure

Figure 6 presents the pairwise Pearson correlation matrix. The three subject scores are mutually highly correlated ($r = 0.96$–0.98), validating the composite as a summary measure. Pupil–teacher ratio is negatively correlated with all three subject scores ($r = -0.58$ to -0.51) and with the composite ($r = -0.558$). GDP per capita is positively correlated with achievement ($r = 0.47$ for the composite). The near-zero correlation between education expenditure share and achievement ($r = -0.095$) stands in sharp contrast to the PTR and GDP correlations.

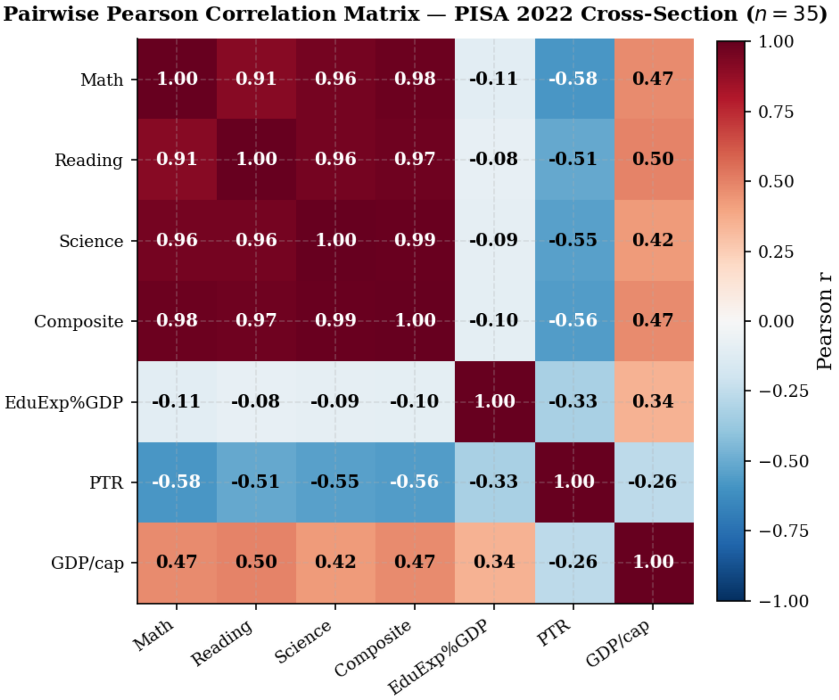

Figure 6: Pairwise Pearson correlation matrix for all seven study variables ($n = 35$). Values in boldface; colour scale runs from blue ($r = -1$) through white ($r = 0$) to red ($r = +1$). PTR = pupil–teacher ratio; EduExp = education expenditure as % GDP; GDP/cap = GDP per capita.

5.7. Model 4: Multiple Regression

Estimating Equation 4 yields:

$$\widehat{\text{PISA}} = 572.1 - 11.0\,\text{EduExp} - 4.28\,\text{PTR} + 0.59\,\text{GDP}, \quad R^2 = 0.592 \tag{6}$$

Table 2 summarises all four models. The multiple regression explains 59.2% of cross-country score variance—a substantial improvement over any single predictor. The pupil–teacher ratio retains a large negative coefficient (-4.28), consistent with Models 1–3. The education expenditure coefficient becomes more negative (-11.0) when PTR and GDP are held constant, reflecting a suppression effect: countries that spend more at a given income level and with a given PTR tend to achieve somewhat less—a pattern consistent with diminishing marginal returns to aggregate investment beyond efficient deployment.

Table 2: Regression results. All models use the 35-country cross-section. Composite PISA score is the dependent variable throughout. SE = OLS standard error.

Variable	Model 1 EduExp	Model 2 PTR	Model 3 GDP	Model 4 Multiple
Intercept	494.8***	538.4***	455.7***	572.1***
EduExp (% GDP)	−2.32	—	—	−11.0
PTR (secondary)	—	−4.097***	—	−4.28***
GDP per capita (k $)	—	—	0.589***	0.595***
R^2	0.009	0.312	0.221	0.592
n	35	35	35	35

*** $p < 0.001$. Model 1 slope is not significant ($p = 0.586$). All models: $n = 35$, OLS.
Model 4 F-statistic significant at $p < 0.001$; SE omitted for space; full output in replication code.

6. Discussion

6.1. The Expenditure Paradox: Why More Spending Does Not Predict Higher Scores

The finding that education expenditure as a share of GDP is uncorrelated with PISA composite scores ($r = -0.095$, $p = 0.586$) is the most policy-relevant result of this analysis. It is consistent with the OECD's own documentation of a non-linear expenditure–outcome relationship: below a cumulative investment threshold of roughly USD 75,000 per student, spending predicts performance; above it, the marginal returns to additional inputs depend almost entirely on how the resources are used rather than how much is spent [1]. Most of the 35 countries in this sample have long surpassed that threshold, which is why aggregate expenditure shares at the country level add little explanatory power for within-sample performance differences.

Singapore is the clearest illustration. It spends only 2.7% of GDP on education—the lowest figure in the sample—yet achieves the highest composite score by a margin of 26 points over second-placed Japan. The explanation lies not in the volume of investment but in its deployment: Singapore maintains comparatively moderate class sizes, invests heavily in teacher selection and professional development, and runs a highly structured curriculum that prioritises deep conceptual understanding in mathematics and science [6]. This pattern is not unique to Singapore: Japan (3.2% GDP, composite 533), Korea (4.6% GDP, composite 523.3), and Estonia (6.6% GDP, composite 515.7) all demonstrate that the returns to education investment are mediated by institutional and pedagogical quality in ways that a single expenditure share indicator cannot capture.

The DEA efficiency analysis by [4] reaches the same conclusion through a different method: their Malmquist productivity estimates show that total factor productivity growth in education is driven primarily by technical progress—new teaching methods, digital resources, assessment innovations—rather than by efficiency catch-up within a given technology set. This implies that the frontier of what is achievable with a given education budget is still moving, and that countries falling behind are not simply insufficiently funded but are failing to adopt effective instructional practices.

6.2. Pupil–Teacher Ratio as the Primary Structural Predictor

The strong negative association between pupil–teacher ratio and composite PISA score ($\hat{\beta} = -4.097$, $R^2 = 0.312$, $p < 0.001$) survives the move to the multiple regression model ($\hat{\beta} = -4.28$, $p < 0.001$), confirming that class size effects are not simply a proxy for income. This result is consistent with the multilevel PISA 2022 analysis by [6], who document that teacher availability and support are among the strongest school-level predictors of mathematics performance across all four high-performing systems they examine—effects that depend on the ratio of teachers to students as a structural precondition.

For middle-income countries in the sample, the PTR finding carries direct practical implications. Mexico (PTR = 28.6, composite = 404.3), Colombia (25.3, 398.7), and Brazil (18.0, 397.3) all have ratios well above the sample mean of 13.7. Bringing Mexico's secondary PTR down to the sample median would, on the basis of the estimated coefficient, predict an increase in composite PISA score of approximately 61 points—equivalent to more than a year and a half of additional learning. This is a large effect and should be interpreted cautiously given the aggregate nature of the estimator; however, it is directionally consistent with quasi-experimental classroom-size research reviewed in [7], which finds persistent positive effects of reduced class sizes on test scores in contexts where baseline ratios are high.

The implication for SDG 4 financing strategy is nuanced. Simply increasing education budgets without directing the additional resources toward teacher hiring and retention is unlikely to produce the achievement gains implicit in SDG 4 target 4.c (substantially increasing the supply of qualified teachers). Conversely, countries that already have low PTRs—Belgium at 8.9, Lithuania at 8.4—do not appear to gain additional achievement advantage from further reductions; their PISA composite scores, while above average, are not at the top of the distribution. This suggests a threshold effect: the marginal return to reducing PTR is highest in the high-ratio range and diminishes as class sizes approach the sample minimum.

6.3. GDP, Overperformance, and the Transferability of Education Reform

GDP per capita explains 22.1% of cross-country PISA variance in the bivariate model, and its coefficient remains virtually unchanged when PTR and education expenditure are added (0.589 versus 0.595), suggesting that the income effect is orthogonal to both instructional structure and spending composition in this sample. The income–achievement correlation reflects the broader socioeconomic environment: parental education levels, cultural attitudes toward schooling, nutrition, housing stability, and out-of-school learning time all co-vary with national income in ways that education budgets alone cannot compensate for [5].

The positive residuals for Estonia (+56.4 points) and Poland (+47.8 points) are therefore particularly instructive. Both countries have restructured their education systems since the early 1990s around comprehensive schooling through age 16, strong teacher training requirements, minimal ability tracking, and curriculum coherence—reforms that were implemented at relatively modest cost given their upper-middle income status at the time of adoption [1]. Their continued outperformance of income predictions in 2022 suggests that these structural reforms have sustained advantages over decades, and that they are genuine policy transfers rather than artefacts of economic or demographic composition.

The large negative residuals for Mexico and Colombia (-44.9 and -37.5 points respectively) indicate the opposite: these economies achieve substantially below what their income level would predict. While both have increased education expenditure as a share of GDP—Colombia at 4.3% is not far from the international benchmark—their PTRs remain among the highest in the sample, and teacher qualification and pay structures limit the depth of instructional capacity that expenditure volumes would otherwise suggest. This combination—adequate aggregate spending shares alongside poor instructional resource deployment—is precisely the configuration that the expenditure paradox in Section 6.1 predicts.

DOI: https://doi.org/10.54216/IJAIET.040203

6.4. Implications for SDG 4 Monitoring and Finance

The 2024 Education Finance Watch estimates that 41 of 80 countries with available data met neither the 4%-of-GDP nor the 15%-of-public-expenditure international benchmarks in 2022 [3]. The present findings suggest that for the majority of high-income countries well above these thresholds, the benchmarks are measuring the wrong thing: they track the volume of investment rather than the efficiency of its deployment. A more useful monitoring framework for high-income OECD contexts would include pupil–teacher ratios, teacher qualification rates, and school autonomy indicators alongside expenditure shares.

For low- and lower-middle-income countries currently below the financing benchmarks, the implication is different: achieving the minimum investment thresholds remains a prerequisite for providing the physical and human infrastructure that learning requires [5, 8]. The SDG 4 financing gap of USD 97 billion annually for the 79 most underfunded countries cannot be bridged by efficiency improvements alone in contexts where basic teacher supply, classroom infrastructure, and materials remain inadequate. For this group, expenditure volume and efficiency are complements rather than substitutes—a distinction that the aggregate result ($r = -0.095$ across all 35 countries) obscures.

6.5. Limitations

Several caveats apply. This is a cross-sectional analysis which cannot determine the direction of causality, that is, the countries with more effective education systems can draw various amounts of public investment as a political economy equilibrium, not that investment leads to effectiveness. The diversity of the 35 country sample omits all of sub-Saharan Africa, South Asia and low income East Asia, the very regions where SDG4 shortfalls are the greatest. In OECD countries, even the PISA database is publicly accessible and can be used to estimate at a level of individuals to enhance the accuracy of all estimates; the most significant extension of the methodology to the results of such studies is a move beyond country-level OLS to student-level hierarchical models. Lastly, the pupil -teacher ratio is a structural measure that is a combination of class size, teacher workload, and subject specialisation in a manner that is not completely addressed by a single national average.

7. Conclusion

This comparison of the results of PISA 2022 in 35 countries reveals three obvious findings. The share of GDP spent on education is not significantly related to composite student achievement in the entire cross-national sample-a pattern that is robust to mathematics, reading, and science in isolated samples and to the inclusion of income and class-size variables. Pupil–teacher ratio: It is a powerful and consistently important negative predictor, and every added student per teacher lowers the composite score by about 4.1 PISA points. These two structural indicators, along with GDP per capita, account 59 percent of cross-country achievement variation, and the pupil teacher coefficient is virtually the same in the bivariate and multiple regression specifications.

The SDG 4 policy message is simple: the question of whether an investment in further education yields learning benefits is not determined by the size of the budgets but the nature of purchases. To high-income systems already beyond the international financing standards, the marginal cost of learning of further expenditure is insignificant unless it is channeled towards the reduction of classes, teacher qualification or instruction time. With lower-income systems yet under the financing levels, sufficient per-student investment must be a condition–but must be coupled with choices of allocation which emphasize the instructional capacity rather than the administrative overhead or infrastructure initiatives with low pedagogical payoff.

The complete analytical code and data description of this article are made publicly accessible in the reproducible version, with only the sources that are free of charge and available in the OECD PISA 2022 database, UNESCO UIS, and World Bank Open Data portals.

References

[1] OECD. PISA 2022 results (Volume I): The State of Learning and Equity in Education. Technical report, OECD Publishing, Paris, France, 2023.

[2] United Nations. Transforming our world: The 2030 agenda for sustainable development, 2015. United Nations General Assembly Resolution A/RES/70/1, New York, NY.

[3] Global Education Monitoring Report Team, UNESCO Institute for Statistics, and World Bank. 2024 education finance watch. Technical report, UNESCO and World Bank, Paris, France and Washington, DC, 2024. URL https://www.unesco.org/gem-report/en/2024efw.

[4] Chaimae Ghernouk and Mariem Liouaeddine. Educational spending efficiency: A comparative analysis using data envelopment analysis and Malmquist index. *Economies*, 14(1):10, 2026. doi: 10.3390/economies14010010.

[5] UNESCO. Global education monitoring report 2024/5: Leadership in education—lead for learning. Technical report, United Nations Educational, Scientific and Cultural Organization, Paris, France, 2024. URL https://www.unesco.org/gem-report/en/2024-leadership.

[6] Jinpeng Niu, Hui Xu, and Jiangtao Yu. Identifying multilevel factors on student mathematics performance for Singapore, South Korea, Finland, and Denmark in PISA 2022: Considering individualistic versus collectivistic cultures. *Humanities and Social Sciences Communications*, 12(1):151, 2025. doi: 10.1057/s41599-025-04466-y.

[7] World Bank. The adequacy of public expenditure on education and the needs post-COVID-19. Technical report, World Bank, Washington, DC, 2023. URL https://thedocs.worldbank.org/en/doc/9b9ecb979e36e80ed50b1f110565f06b-0200022023/original/Adequacy-Paper-Final.pdf.

[8] Zhiying Zhao and Zhanglin Sidi. Exploring the impact of education on sustainable resource efficiency: The interplay with CO_2 emissions, renewable energy, and agriculture in G20 nations. *Frontiers in Environmental Science*, 13:1623978, 2025. doi: 10.3389/fenvs.2025.1623978.

[9] UNESCO Institute for Statistics. UIS education statistics: Government expenditure on education as a percentage of GDP. Data portal, 2024. URL `https://uis.unesco.org/bdds`.

[10] World Bank. World development indicators. Technical report, World Bank, Washington, DC, 2024. URL `https://databank.worldbank.org/source/world-development-indicators`.