Int. J. of AI and Education Technology

♦ ISSN: 2835-2432 Vol: 04, No. 02, 2025 ♦ PP. 17–25

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



FULL LENGTH ARTICLE

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, Lebanese International University, Lebanon

Emails: reshmaarhaan2010@gmail.com; Hanadi.diab@liu.edu.lb

Received: April 03, 2025 Revised: May 04, 2025 Accepted: July 03, 2025 ⋆ Corresponding author

ABSTRACT

All people must have access to educational opportunities that 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 countries achieve different results from their spending levels. This study uses country-level PISA 2022 scores, the first international assessment data published after COVID-19, to analyze public education expenditure as a share of GDP together with the pupil–teacher ratio and GDP per capita in relation to student academic performance across three subjects. Public education funding as a percentage of GDP is not statistically associated with PISA scores across 35 countries ($r = -0.095$, $p = 0.586$). The pupil–teacher ratio is a strong negative predictor of student performance ($\hat{\beta} = -4.097$, $R^2 = 0.312$, $p < 0.001$). A three-variable regression model combining expenditure share, 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 systems, notably Estonia and Poland, substantially outperform their GDP-predicted scores. The results indicate that education systems should focus resources on strengthening instructional capacity.

Keywords: PISA 2022 ▪ Education expenditure ▪ Pupil–teacher ratio ▪ Student achievement ▪ Sustainable Development Goal 4 ▪ Cross-national analysis ▪ Education finance

1. INTRODUCTION

In December 2023, the OECD Programme for International Student Assessment (PISA) sent a clear message: among the 37 OECD countries with PISA 2022 results, mean mathematics performance decreased by a record 15 points and reading decreased by 10 points, the largest consecutive decline in the history of the programme since its inception in 2000 [1]. The decline in reading predates COVID-19, but mathematics appears to capture both a long-term systemic trend and a crisis shock in a single assessment, following pandemic-related disruptions [1]. These findings rekindle debates about student learning: when investment does not produce commensurate results, what explains differences in achievement across nations?

This question lies at the core of Sustainable Development Goal 4 (SDG 4), which aims to provide inclusive, equitable, and quality education for everyone by 2030 [2]. Progress on SDG 4 indicators has lagged across many aspects. In 2024, median public education spending worldwide was 4.2% of GDP, down from 4.3% in 2010, and only 22% of countries

© 2025 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

had simultaneously reached the international financing targets of 4% of GDP and 15% of government spending [3]. However, the relationship between spending and learning is not linear: OECD data indicate that below cumulative expenditure of USD 75,000 per student during primary and secondary education, higher investment is associated with higher PISA scores, but beyond that level the association weakens [1]. Understanding success in high-spending systems is therefore as important as addressing investment gaps in low-income environments.

This paper contributes to this question in three ways. It provides a repeatable cross-national study of PISA 2022 based on publicly available institutional data. It isolates the associations of aggregate expenditure, instructional resource intensity (pupil–teacher ratio), and economic development with achievement, showing that these inputs have qualitatively different relationships with student performance. Finally, it places the findings in the SDG 4 accountability framework to inform the prioritization of education investments by ministries and development banks.

The paper is organized as follows. Section 2 reviews the empirical literature. Section 3 presents the conceptual framework and hypotheses. Section 4 describes the data and methodology. Section 5 reports the results, Section 6 discusses the findings, and Section 7 concludes.

2. LITERATURE REVIEW

2.1 Education Expenditure and Learning Outcomes

Whether increased education spending yields better student outcomes has occupied education economists for decades, and the answer is not simple. Quasi-experimental school-finance reform studies in national settings generally report small positive returns: a sustained increase in per-pupil spending of USD 1,000 per year over 10 years is associated with a 0.12–0.24 standard-deviation improvement in test scores in the United States [1]. The association is weaker internationally. Cross-national studies find that the aggregate spending share of GDP is a poor predictor of PISA performance after controlling for income [1]. High-income countries can spend more while also benefiting from institutional and human-capital environments that support learning, irrespective of the marginal dollar spent on school budgets.

OECD's Education at a Glance 2024 presents comparative indicators on education finance alongside student outcomes, teacher resources, learning environments, and school organization. This broader framework supports analyzing funding together with instructional and socioeconomic factors rather than treating the spending share as a direct measure of efficiency [4].

2.2 Instructional Resources: Teacher Quality and Pupil–Teacher Ratio

If aggregate spending 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 teacher qualifications and pupil–teacher ratios have more direct associations with student performance than aggregate expenditure totals [5]. OECD analyses of PISA 2022 show that intrinsic motivation predicts greater uptake of learning strategies, and students who regularly use these strategies tend to outperform peers in mathematics even after socioeconomic adjustment [6]. This evidence highlights student engagement as an important learning process that aggregate expenditure totals do not capture. Cross-national evidence on pupil–teacher ratios is consistent: smaller secondary classes create conditions for more individualized instruction, and their positive association with achievement holds across diverse income settings [7].

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

At the country level, GDP per capita is the best single predictor of PISA performance and serves in large part as a proxy for the wider socioeconomic context in which schooling takes place [1]. In 2022, socioeconomically advantaged students scored an average of 93 points higher than disadvantaged students across OECD countries, a difference equivalent to about three years of learning. This average masks substantial variation: in Macao (China), disadvantaged students scored above the OECD average, demonstrating that high system-wide equity can be achieved without high GDP per capita. Estonia and Poland have also consistently outperformed their income-predicted results; their experience with teacher professionalism, later ability tracking, and high curriculum standards is increasingly proposed as a model for middle-income reform [1].

The SDG 4 financing literature indicates that narrowing the global learning gap requires increased investment in underfunded low-income environments and more efficient allocation of existing resources in middle- and high-income systems [3]. The annual financing gap for achieving SDG 4 targets was estimated at USD 97 billion across 79 low- and lower-middle-income countries in 2023, with the largest shortfall in sub-Saharan Africa [5]. Evidence about which investments yield the greatest returns per dollar is therefore central to the empirical question addressed here.

2.4 Education, Sustainability, and the SDG 4 Nexus

Education is embedded in wider institutional and social settings that shape progress toward SDG 4. The 2024/5 Global Education Monitoring Report identifies leadership at school, system, and political levels as an important influence on inclusive, equitable, quality education, emphasizing the role of organizational and contextual conditions [5]. This perspective supports examining expenditure alongside instructional resources and socioeconomic context in cross-national analyses.

3. CONCEPTUAL FRAMEWORK AND HYPOTHESES

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

The instructional-input channel (H1) links expenditure to the deployment of teaching staff: higher spending, when efficiently allocated, should reduce class sizes and improve teacher qualifications, with a direct route to student outcomes. The governance and institutional channel (H2) captures how

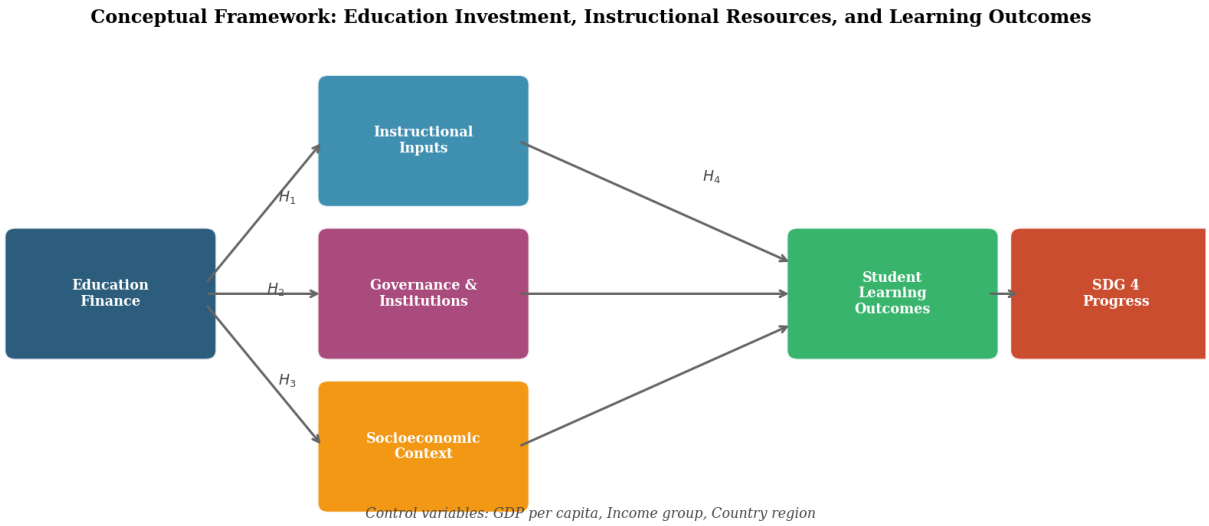


Figure 1. Conceptual framework linking public education finance to student learning outcomes through instructional inputs (H1), institutional governance (H2), and socioeconomic context (H3). Together, these pathways are hypothesized to explain cross-national variation in PISA achievement (H4). Control variables are listed at the base of the diagram.

system-level decisions—including curriculum design, assessment policy, teacher evaluation, and school autonomy—translate financial inputs into pedagogical practice. The socioeconomic-context channel (H3) acknowledges that learning outcomes reflect not only in-school resources but also the broader material and cultural environment in which students learn.

Four testable hypotheses follow. H1: countries with lower pupil–teacher ratios achieve higher mean PISA scores. H2: public education expenditure as a share of GDP is positively associated with composite PISA scores, net of income effects. H3: GDP per capita is positively associated with composite PISA scores. H4: 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 based on data completeness for the four study variables. The final set spans high-income OECD members and several upper- and lower-middle-income participants, including Chile, Turkey, Mexico, Colombia, and Brazil, providing cross-income variation for testing H2 and H3.

Four variables are employed. PISA 2022 mathematics, reading, and science scores, and their composite mean, are drawn from Table I.B1.1 of the OECD PISA 2022 Results, Volume I, released on December 5, 2023 [1]. The public PISA 2022 dataset, including microdata, 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," using the most recent available year for each country (predominantly 2019–2021) [8]. The secondary-education pupil–teacher ratio is from World Bank World Development Indicators (WDI), series SE.SEC.STUDT.TCH.ZS [9]. GDP per capita (thousand current USD, 2022) is from World Bank WDI series NY.GDP.PCAP.CD [9].

Table 1. Descriptive statistics for the cross-sectional sample ($n = 35$).

Variable	Mean	SD	Min	P25	Median	Max
Mathematics score (points)	476.5	38.5	379	470	482	575
Reading score (points)	481.3	28.4	409	474	480	543
Science score (points)	489.4	34.3	403	477	492	561
Composite PISA score (points)	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 (thousand 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.

4.2 Descriptive Statistics

Table 1 reports summary statistics for all study variables. The mean composite PISA score across the 35 countries is 482.4 points (SD = 33.1), ranging 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 ranges from USD 6,900 (Colombia) to USD 106,600 (Norway).

4.3 Model Specification

Three bivariate ordinary least-squares (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, or 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 using the t-statistic $t = \hat{\beta}_k / \widehat{\sigma}_{\hat{\beta}_k}$ 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 $\hat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$. Pairwise Pearson correlations are computed over the full sample using:

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

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 between 485 and 492. At the lower end, Brazil (397.3) and Colombia (398.7) are well below the OECD mathematics average of 472. A clear stratification between high-income and other economies is visible, although Estonia and Poland score above their income band, and Turkey approaches the European mid-range despite much lower GDP, suggesting that system factors beyond income matter considerably.

Panel (b) reveals no 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%, and Finland at 6.8%—achieve high but not disproportionate scores, while Singapore has the highest score in the dataset and 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 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 ratios—Belgium (8.9), Lithuania (8.4), Slovenia (8.7), and Austria (9.9)—score 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) shows the expected positive income–achievement gradient but also 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 has 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 less influence from top outliers. 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.

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:

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

The slope is negative and far from statistically significant ($t = -0.55$). The Pearson correlation is $r = -0.095$, confirming the absence of a meaningful linear association. Disaggregation 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 supports the OECD finding that, within the high-expenditure tier, outcomes are less predictive of investment level than of the efficiency and composition of resource deployment [1].

Figure 4 examines this relationship further. Panel (a) ranks countries by expenditure share, distinguishing high spenders ($\geq 5.5\%$ of GDP, $n = 18$) from lower spenders ($n = 17$). Panel (b) compares mean subject scores between these groups and the OECD average. High spenders achieve a mean composite score of 479.4 versus 485.6 for lower spenders, a difference of 6.2 points in the direction opposite to the investment hypothesis. Neither group differs significantly from the other or from the OECD average. The directional reversal suggests a compositional effect: the high-spending group includes Scandinavian countries alongside Chile and Brazil, while the low-spending group includes Singapore and Japan, whose lower expenditure coexists with top-tier outcomes.

5.4 Model 2: Pupil–Teacher Ratio and Achievement

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

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

The slope is negative and highly significant ($t = -3.87$, $p = 0.0005$), indicating that each additional student per teacher is

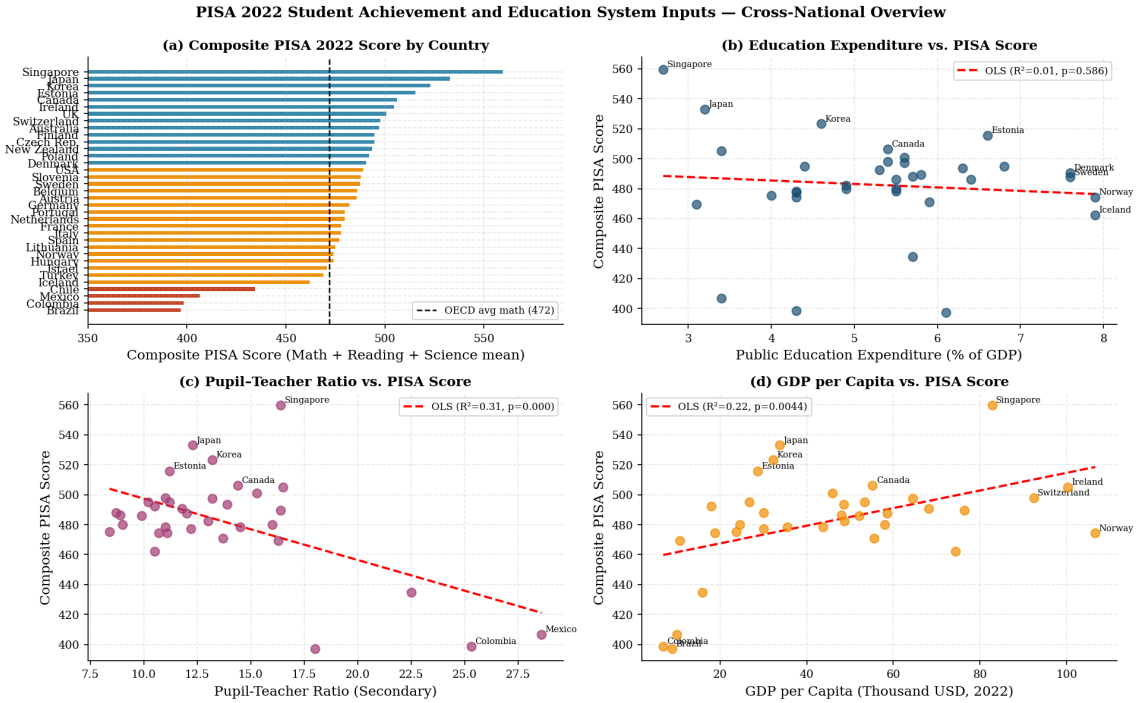


Figure 2. Four-panel overview of PISA 2022 outcomes and education-system inputs. (a) Countries ranked by composite PISA score (colour bands: blue ≥ 490, orange 450–489, red < 450); (b) education expenditure (% GDP) versus composite score with OLS fit; (c) pupil–teacher ratio versus composite score with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

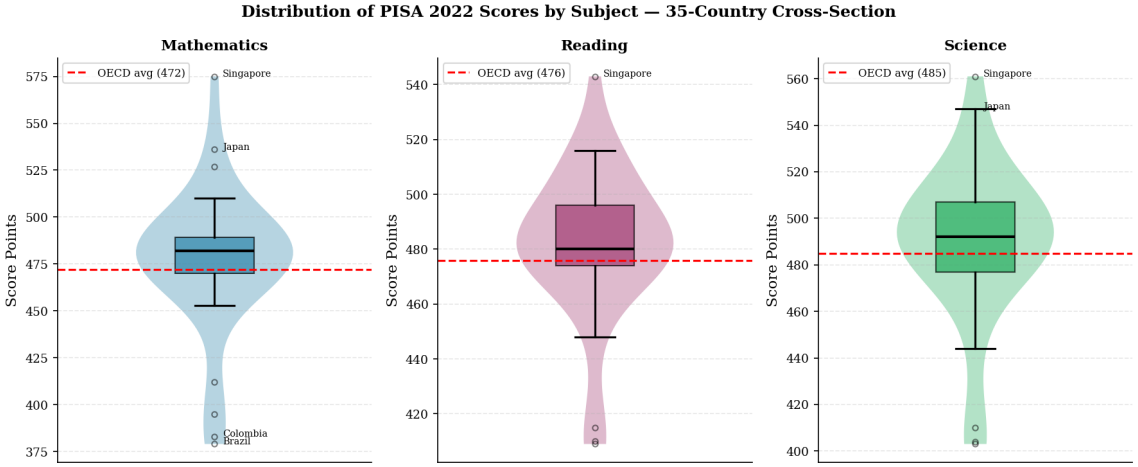


Figure 3. Violin and box plots of PISA 2022 score distributions for mathematics, reading, and science across 35 countries. Red dashed lines mark OECD averages (mathematics: 472; reading: 476; science: 485). Labelled points are notable high or low performers.

associated with a reduction of approximately 4.1 PISA points in the composite score. Applying this coefficient 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 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:

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

The slope is positive and significant ($t = 3.06$), consistent with H3. 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, substantially outperforming their income predictions. Mexico (−44.9), Colombia (−37.5), and Brazil (−30.2) are the largest negative outliers, achieving below what their income level would predict.

Panel (b) shows composite-score distributions by income group. High-income economies (above USD 35,000 per capita) have a median composite of approximately 489 points, well above the OECD mathematics average. Upper-middle-income countries (USD 15,000–35,000) show wider variance, with Estonia and Poland at the high end and the Czech Republic and Slovenia near the median. Lower-income economies (below USD 15,000) in the sample—Colombia, Brazil, Mexico,

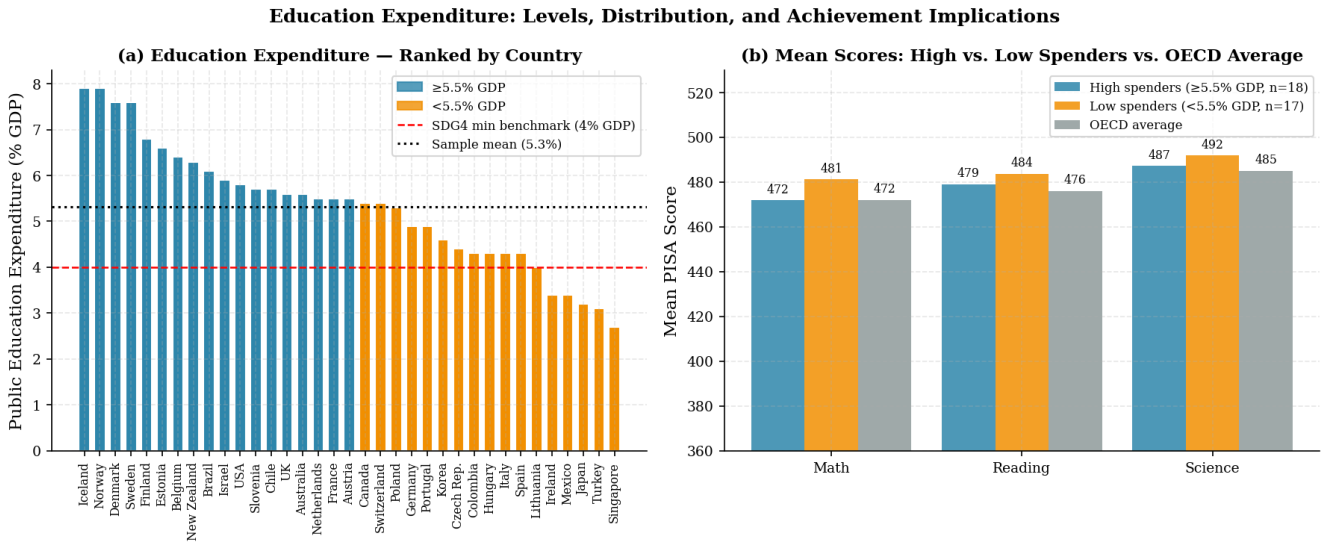


Figure 4. Education-expenditure analysis. (a) Countries ranked by education-expenditure share of GDP; blue bars indicate high spenders ($\geq 5.5\%$ GDP), orange bars indicate lower spenders. The red dashed line marks the SDG 4 minimum benchmark of 4% of GDP. (b) Mean subject scores for high spenders,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

and Turkey—cluster near the bottom. Panels (c) and (d) replot expenditure and pupil–teacher-ratio scatter plots with income groups colour-coded. 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. The pupil–teacher-ratio relationship holds within each income group, suggesting that class size is an independent predictor beyond income.

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 ($r = -0.095$) contrasts with the pupil–teacher ratio and GDP correlations.

5.7 Model 4: Multiple Regression

Estimating Equation 4 yields:

$$\begin{aligned} \text{PISA} &= 572.1 - 11.0\,\text{EduExp} - 4.28\,\text{PTR} + 0.59\,\text{GDP}, \\ R^2 &= 0.592. \end{aligned} \tag{6}$$

Table 2 summarizes 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 for the 35-country cross-section. Composite PISA score is the dependent variable throughout.

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 (thousand USD)	—	—	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 use OLS; standard errors are omitted for space. The Model 4 F-statistic is significant at $p < 0.001$; full output is available in the 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 OECD documentation of a nonlinear expenditure–outcome relationship: below a cumulative investment threshold of roughly USD 75,000 per student, spending predicts performance; above it, marginal returns depend almost entirely on how resources are used rather than how much is spent [1]. Most countries in this sample have surpassed that threshold, which helps explain why aggregate expenditure shares 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. This contrast shows that education expenditure share alone does not capture the learning processes associated with achievement. OECD analysis of PISA 2022 shows that intrinsic motivation predicts students' uptake of learning strategies, and students who use these strategies regularly tend to outperform peers in mathematics even after socioeconomic adjustment, pointing to learning processes not

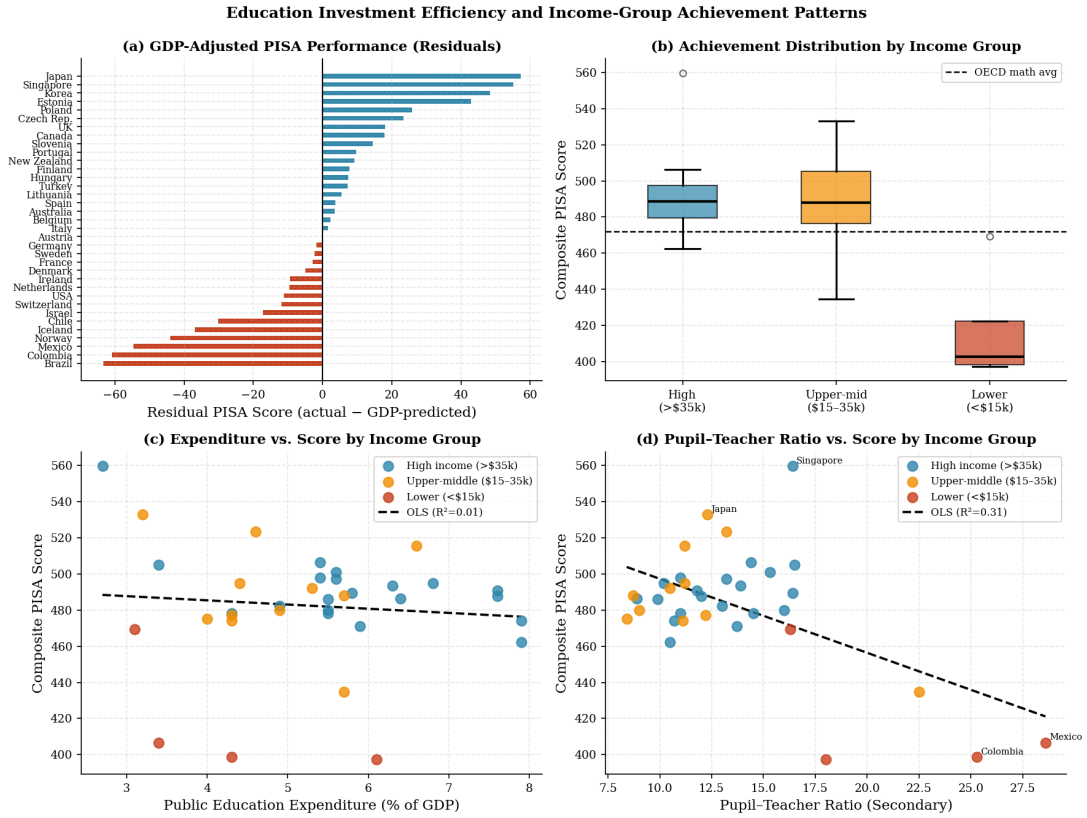


Figure 5. Income-adjusted achievement patterns. (a) Residuals from Model 3 (actual minus GDP-predicted composite score); positive values indicate overperformance relative to income. (b) Composite-score distribution by income group (box plots). (c) Education expenditure versus score, coloured by income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

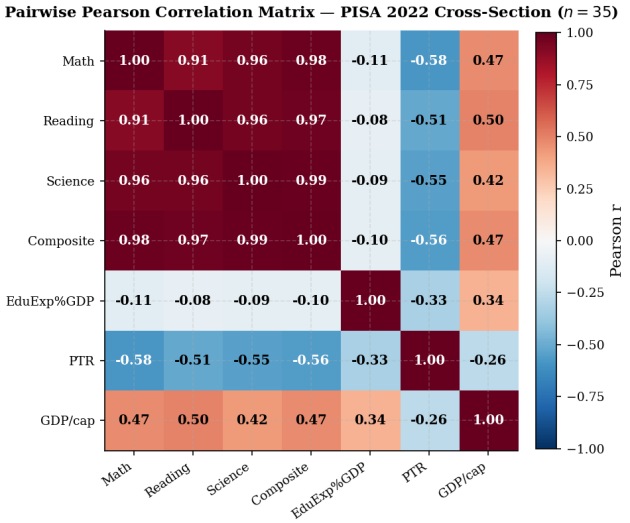


Figure 6. Pairwise Pearson correlation matrix for all seven study variables ($n = 35$). Values are shown in bold; the colour scale runs from blue ($r = -1$) through white ($r = 0$) to red ($r = +1$). PTR = pupil–teacher ratio; EduExp = education expenditure as % GDP; GDP/cap = GDP per capita.

reflected in budget shares [6]. Japan (3.2% GDP, composite 533), Korea (4.6% GDP, composite 523.3), and Estonia (6.6% GDP, composite 515.7) likewise show that high achievement can coincide with varied expenditure shares.

OECD's Education at a Glance 2024 presents expenditure indicators alongside measures of education outcomes, teachers, learning environments, and the organization of education systems. This comparative perspective reinforces the need to interpret spending levels alongside system context; aggregate expenditure shares alone cannot identify efficiency or technical progress [4].

6.2 Pupil–Teacher Ratio as the Primary Structural Predictor

The strong negative association between pupil–teacher ratio and composite PISA score ($\hat{\beta} = -4.097$, $R^2 = 0.312$, $p < 0.001$) persists in the multiple-regression model ($\hat{\beta} = -4.28$, $p < 0.001$), indicating that the PTR association is not simply a proxy for income.

For middle-income countries in the sample, the pupil–teacher-ratio finding has 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. On the basis of the estimated coefficient, bringing Mexico's secondary PTR down to the sample median would predict an increase in composite PISA score of approximately 61 points, equivalent to more than a year and a half of additional learning. This large effect should be interpreted cautiously given the aggregate nature of the estimator; however, it is directionally consistent with quasi-experimental classroom-size research reviewed by the World Bank, which finds persistent positive effects of reduced class sizes where baseline ratios are high [7].

The implication for SDG 4 financing strategy is nuanced. Increasing education budgets without directing additional resources toward teacher hiring and retention is unlikely to produce the achievement gains implicit in SDG 4 target 4.c, which calls for substantially increasing the supply of qualified teachers. Conversely, countries that already have low PTRs—Belgium at 8.9 and Lithuania at 8.4—do not appear to gain additional achievement advantage from further reductions;

their composite PISA 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). This suggests that the income effect is orthogonal to instructional structure and spending composition in this sample. The income–achievement correlation reflects the broader socioeconomic environment: parental education, cultural attitudes toward schooling, nutrition, housing stability, and out-of-school learning time all covary 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 particularly instructive. Both countries restructured their education systems since the early 1990s around comprehensive schooling through age 16, strong teacher-training requirements, minimal ability tracking, and curriculum coherence. These reforms were implemented at relatively modest cost given their upper-middle-income status at the time [1]. Their continued outperformance of income predictions in 2022 suggests that these structural reforms have sustained advantages and may be transferable policy models.

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. Although 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 instructional capacity that expenditure volumes might otherwise suggest. This combination of adequate aggregate spending shares and weak instructional-resource deployment is precisely the expenditure paradox described in Section 6.1.

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 most high-income countries already above these thresholds, the benchmarks measure investment volume 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, achieving minimum investment thresholds remains a prerequisite for providing the physical and human infrastructure that learning requires [3, 5]. The annual SDG 4 financing gap of USD 97 billion for the 79 most underfunded countries cannot be bridged by efficiency improvements alone where basic teacher supply, classroom infrastructure, and materials remain inadequate. For this group, expenditure volume and efficiency are complements rather than substitutes, a distinction obscured by the aggregate result ($r = -0.095$ across all 35 countries).

6.5 Limitations

Several caveats apply. This cross-sectional analysis cannot determine the direction of causality: countries with more effective education systems may draw varying amounts of public investment as a political-economy equilibrium, rather than investment necessarily producing effectiveness. The sample of 35 countries omits all of sub-Saharan Africa, South Asia, and low-income East Asia, where SDG 4 shortfalls are greatest. The publicly accessible PISA database could support individual-level estimation in OECD countries and improve the accuracy of estimates; a major methodological extension would move beyond country-level OLS to student-level hierarchical models. Finally, the pupil–teacher ratio is a structural measure combining class size, teacher workload, and subject specialization that cannot be fully represented by a single national average.

7. CONCLUSION

This comparison of PISA 2022 results in 35 countries yields three main findings. The share of GDP spent on education is not significantly related to composite student achievement in the cross-national sample, a pattern that is robust across mathematics, reading, and science and to the inclusion of income and class-size variables. The pupil–teacher ratio is a powerful and consistently important negative predictor: each additional student per teacher is associated with about 4.1 fewer PISA points in the composite score. These structural indicators, together with GDP per capita, account for 59% of cross-country achievement variation, and the pupil–teacher-ratio coefficient is nearly unchanged between the bivariate and multiple-regression specifications.

The SDG 4 policy message is that the learning benefits of further education investment depend not only on budget size but also on how resources are allocated. For high-income systems already beyond the international financing standards, the marginal learning return to further expenditure is limited unless resources are directed toward smaller classes, teacher qualifications, or instructional time. For lower-income systems still below financing thresholds, sufficient per-student investment remains necessary, but it should be coupled with allocation choices that prioritize instructional capacity over administrative overhead or infrastructure with low pedagogical payoff.

The complete analytical code and data description are made publicly accessible in the reproducible version of this article, using freely available sources from 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,” OECD Publishing, Paris, France, Tech. Rep., 2023.

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

[3] Global Education Monitoring Report Team and UNESCO Institute for Statistics and World Bank, “2024 education finance watch,” UNESCO and World Bank, Paris, France and Washington, DC, Tech. Rep., 2024.

[4] OECD, “Education at a glance 2024: OECD indicators,” OECD Publishing, Paris, France, Tech. Rep., 2024.

[5] UNESCO, “Global education monitoring report 2024/5: Leadership in education—lead for learning,” United Nations Educational, Scientific and Cultural Organization, Paris, France, Tech. Rep., 2024.

[6] OECD, “PISA 2022 results (volume v): Learning strategies and attitudes for life,” OECD Publishing, Paris, France, Tech. Rep., 2024.

[7] World Bank, “The adequacy of public expenditure on education and the needs post-covid-19,” World Bank, Washington, DC, Tech. Rep., 2023.

[8] UNESCO Institute for Statistics, “UIS education statistics: Government expenditure on education as a percentage of GDP,” UNESCO Institute for Statistics, Tech. Rep., 2024.

[9] World Bank, “World development indicators,” World Bank, Washington, DC, Tech. Rep., 2024.