Socioeconomic Gradients in Student Achievement: A Cross-National Decomposition Using PISA 2022

David Goh
daveed@cs.toronto.edu
University of Toronto
Toronto, Canada

ACM Reference Format:
David Goh. 2026. Socioeconomic Gradients in Student Achievement: A Cross-National Decomposition Using PISA 2022. In . N/A, 3 pages. https://doi.org/N/A

1 Introduction

Socioeconomic status (SES) remains the most significant predictor of student academic outcomes across educational systems internationally. Prior research consistently shows that family background—measured by parental education, occupational status, and income—is strongly associated with student performance in core subjects like reading, mathematics, and science. However, current cross-national studies, such as those relying on the PISA Index of Economic, Social, and Cultural Status (ESCS), typically use a single composite index. This approach limits the ability to isolate and understand the distinct contributions of individual SES components.

This study addresses this gap by explicitly decomposing SES into three distinct components: parental education, parental occupational status, and household wealth/resources (proxied by ESCS). We examine their relative influence on student achievement across countries participating in PISA 2022, applying strict data quality criteria for methodological robustness.

This work extends existing research in three critical ways. First, through SES Decomposition, we move beyond composite measures to compare the separate effects of parental education, occupational status, and resources on performance. Second, by analyzing Cross-Country Patterns, we estimate multi-component SES–achievement gradients to highlight international variation in the steepness of these disparities, identifying contexts with unusually high or low SES effects. Finally, through Subject-Specific Analysis, we explore whether the influence of SES components differs across reading, mathematics, and science domains. This descriptive and exploratory analysis does not make causal claims but provides a nuanced account intended to inform educational policy by clarifying which SES dimensions drive achievement disparities.

2552, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/N/A

2 Review of Literature

Understanding these findings requires integration with three key areas of international literature.

2.1 Income–Achievement Gradients

Chmielewski and Reardon (2016) [1] provided a foundational perspective on cross-national income–achievement gaps, demonstrating that gaps measured by direct family income differ substantively from those measured by composite SES indices like ESCS. They highlighted the challenge of accurately capturing income-based disparities. Our study builds on this by employing a decomposition approach, breaking the ESCS index into its constituent parts to enable a more granular assessment of which specific elements—education, occupation, or wealth—drive cross-country differences.

2.2 SES–Achievement Slopes and Subject-Specific Patterns

Jacobs and Wolbers (2018) [2] analyzed the influence of the composite ESCS index on the probability of top performance in PISA 2015 reading and mathematics. They observed modest subject-specific differences in the SES gradients, suggesting that the link between background and achievement may vary by domain. While their use of a composite measure limited insight into individual components, their findings underscore the necessity of our subject-specific focus, which estimates separate gradients for reading, mathematics, and science while decomposing the SES factors.

2.3 System-Level Context: Tracking and Inequality

The role of systemic features is a crucial context for SES disparities. Terrin and Triventi's (2022) [3] meta-analysis concluded that educational tracking modestly increases inequality without significantly impacting overall achievement. However, this meta-analysis was largely based on studies that treated SES as a single factor. By decomposing SES, our study provides clearer insight into which specific dimensions of socioeconomic background are most sensitive to educational system characteristics, such as the age at which students are tracked into different educational paths.

3 Methodology

3.1 Data Source and Sample

This study uses data from all countries participating in PISA 2022, including OECD and non-OECD members. Analyses were restricted to students who completed at least six years of formal schooling. Countries were included if they had at least 5,000 students with valid SES data and less than 10

Applying these strict inclusion criteria drastically reduced the number of countries in the analysis. While this limits the breadth of cross-national comparisons, it ensures that the estimated SES–achievement gradients are based on reliable and comparable data. This rigorous approach prioritizes methodological validity over sample size, highlighting the importance of careful data curation in international assessments.

3.2 Outcome Variables

Academic achievement was measured in three core domains: reading, mathematics, and science. PISA provides ten plausible values (PVs) per domain to account for measurement uncertainty due to matrix sampling. All analyses were conducted separately for each PV, with results combined using Rubin's rules to produce final population-level estimates.

3.3 SES Decomposition

SES was decomposed into three components:

- Parental Education (PAREDINT): Highest parental attainment, standardized across countries.
- Parental Occupational Status (HISEI): Standardized occupational index.
- Household Wealth/Resources (ESCS): Composite index including parental occupation, educational resources at home, and cultural possessions.

Each component was analyzed independently, enabling direct comparison of effect sizes on student achievement.

3.4 Estimation of SES–Achievement Gradients

For each country, subject, and SES component, we fitted weighted least squares regressions of student achievement on the SES component, using student weights (`W_FSTUWT`). Regressions were run separately for each PV. For example, for subject j and SES component k in country i:

$$Y_{ij}^{(pv)} = \beta_0 + \beta_1 \mathrm{SES}_{ik} + \epsilon_{ij} \tag{1}$$

where $Y_{ij}^{(pv)}$ is the student's score on PV pv, and β_1 is the SES gradient. Coefficients and standard errors across the ten PVs were combined using Rubin's rules:

$$\bar{\beta} = \frac{1}{m} \sum_{pv} \beta^{(pv)}, \qquad T = \bar{U} + \left(1 + \frac{1}{m}\right) B \tag{2}$$

where $\bar{U}$ is the average within-PV variance, B is the between-PV variance, and $m = 10$ is the number of PVs. The square root of T gives the total standard error.

3.5 Cross-Country Comparisons and Normalization

Country-level SES gradients were derived from regression coefficients. For visual comparisons, coefficients for each SES component were normalized per country by dividing each component by the sum of the three SES coefficients, allowing proportionate comparisons across countries. Countries with extreme gradients were highlighted descriptively.

3.6 Subject-Specific Analyses

Gradients were estimated separately for reading, mathematics, and science. Modeling choices—including weights, PV handling, and SES decomposition—were applied consistently across subjects. Differences in effect sizes were interpreted descriptively and summarized using mean and standard deviation across countries.

3.7 Linking SES Gradients to Structural Features

To examine the potential role of system-level factors, we merged country-level ESCS gradients with the earliest age at which students were tracked into differentiated educational pathways. Analyses included boxplots, scatterplots with regression lines, and aggregated summaries (mean and standard deviation of ESCS, PAREDINT, and HISEI coefficients across countries).

3.8 Visualization

Normalized SES contributions were visualized using stacked bar plots and heatmaps for cross-country comparisons, and scatterplots and boxplots for tracking-age analyses. These visualizations complement numerical summaries, communicating both country-level variation and subject-specific patterns.

4 Results

4.1 SES Component Contributions Across Countries

We examined the contributions of household wealth/resources (ESCS), parental education (PAREDINT), and parental occupational status (HISEI) to PISA 2022 achievement gradients. As illustrated by normalized coefficient plots (Figure 1), ESCS represents the largest component of the SES gradient in all countries included in the study.

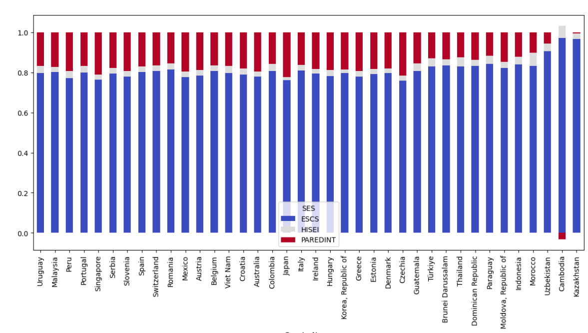

Figure 1: Proportional Contributions of Decomposed Socioeconomic Status (SES) Components to PISA 2022 Achievement Gradients. (Normalized by country, showing the relative weight of Household Resources (ESCS), Parental Education (PAREDINT), and Parental Occupational Status (HISEI) in driving the overall SES effect.)

Despite the dominance of household resources, the relative contribution of parental education (PAREDINT) varied substantially. Countries like Japan, Czechia, and Singapore showed higher PAREDINT contributions, indicating a more substantial independent role for parental education in these systems. Conversely, Cambodia presented a unique pattern where a dominant ESCS contribution was accompanied by a negative PAREDINT coefficient, an anomaly suggesting that, once resources are accounted for, the association between parental education and achievement reverses, potentially due to context-specific schooling access or measurement issues. Parental occupational status (HISEI) played a minimal role across most countries once the other two components were included, suggesting significant overlap with education and household resources.

4.2 Subject-Specific Patterns in SES Components

To examine domain-specific effects, we calculated mean coefficients for ESCS, PAREDINT, and HISEI separately for mathematics, reading, and science. The contribution of each SES component to student achievement was nearly identical across all subjects (Table 2). Mean coefficients showed minimal variation between domains, with differences well within one standard deviation.

Subject	MATH	READ	SCIE
SES			
ESCS	32.81 ± 12.48	32.63 ± 11.76	32.34 ± 12.74
HISEI	1.20 ± 0.39	1.21 ± 0.39	1.19 ± 0.41
PAREDINT	7.19 ± 3.99	7.09 ± 3.72	7.03 ± 3.86

Figure 2: Table: Mean coefficients (± standard deviation) for SES components across countries, by subject. Values represent the average strength of association between each SES component and achievement in mathematics, reading, and science. The consistency across subjects indicates that SES components operate similarly regardless of academic domain.

4.3 SES Gradients and Age of Tracking

Boxplots of ESCS coefficients by tracking age (Figure 3) revealed that countries with earlier tracking (ages 10–12) tend to show higher median SES gradients compared to those with later tracking (ages 15–16). However, the distributions overlap substantially, indicating significant heterogeneity.

Faceted scatterplots confirmed a negative association between tracking age and the ESCS gradient across all three subjects, with Science exhibiting the steepest gradient, followed by mathematics and reading.

5 Interpretation and Contribution

These analyses complement our SES decomposition by showing that policy-relevant contextual factors—here, the age of tracking—coincide with differences in how SES translates into achievement. By linking ESCS gradients to structural features of educational systems, we provide descriptive evidence that contextual factors may amplify or mitigate the role of socioeconomic background. Combined with our finding that household wealth and resources (ESCS) dominate over parental education and occupational status in most countries, these results help interpret cross-country variation in SES effects and suggest that both within-family resources and system-level structures contribute to observed achievement disparities.

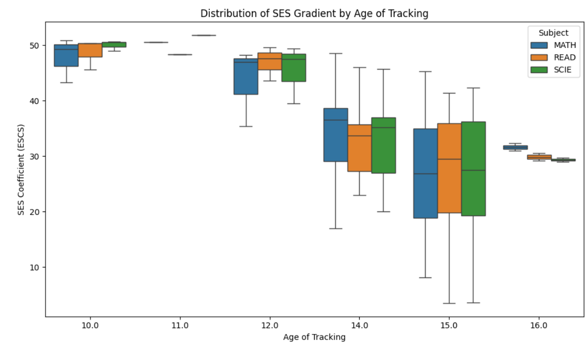

Figure 3: Distribution of ESCS coefficients by age of tracking, faceted by subject. Each boxplot shows the spread of SES gradient strengths for countries that track students at a given age. Wide distributions at ages 14 and 15 reflect substantial heterogeneity among countries with later tracking systems.

The consistency of SES component effects across subjects, coupled with the variation in gradient strength related to tracking age, indicates that while the fundamental mechanisms linking socioeconomic advantage to achievement operate similarly across domains, structural features of educational systems shape the magnitude of these relationships. Future research examining causal mechanisms and additional institutional features will build on this descriptive foundation to clarify how educational policies can most effectively reduce socioeconomic disparities in student achievement.

Acknowledgments

This research was prepared by David Goh. Email: `daveed@cs.toronto.edu`. The author thanks colleagues who provided feedback and acknowledges the PISA data providers. This work uses public microdata from PISA 2022 and makes descriptive cross-country comparisons only.

References

[1] Adam K. Chmielewski and Sean F. Reardon. 2016. Patterns of Cross-National Variation in the Association Between Income and Academic Achievement. *AERA Open* 2, 3 (2016), 2332858416649593. doi:10.1177/2332858416649593

[2] Bart Jacobs and Maarten H. J. Wolbers. 2018. Inequality in top performance: An examination of cross-country variation in excellence gaps across different levels of parental socioeconomic status. *Educational Research and Evaluation* 24, 1-2 (2018), 68–87. doi:10.1080/13803611.2018.1520130

[3] Émilie Terrin and Moris Triventi. 2023. The Effect of School Tracking on Student Achievement and Inequality: A Meta-Analysis. *Review of Educational Research* 93, 2 (2023), 236–274. doi:10.3102/00346543221100850