IUSSP 2025 – Extended Abstract

REGIONAL DISPARITIES IN EDUCATION AND SCHOOL ENVIRONMENT QUALITY IN BRAZIL: A DATA-DRIVEN ANALYSIS

INTRODUCTION

Human capital formation is closely linked to education quality and the school environment. Schools with adequate infrastructure and appropriate resources provide more meaningful learning, contributing to the development of essential skills for economic and social progress[1]. The concept of "healthy buildings"[2] reinforces this connection, as well-structured and healthy school environment promote well-being, enhance academic performance and foster human capital development.

The aim of this article is to investigate, through a comprehensive, data-driven approach, the relationship between quality of education and the quality of the school environment. The analysis focuses on regional disparities among Brazilian municipalities and changes observed over time.

The quality of public education in Brazil is analyzed from two correlated perspectives. The first refers to the school environment's quality, defined by a set of characteristics related to infrastructure, equipment, and available resources. For this, we propose a novel, robust indicator that synthesizes school quality in a multidimensional way. The uniqueness of this indicator lies in its scope: it is a longitudinal, census-based, and annual measure, applied to all Brazilian schools, both public and private, regardless of size. These characteristics provide a detailed and continuous assessment of school environment quality over time.

The second perspective addresses the quality of education offered by schools, measured by the Basic Education Development Index (IDEB), which combines data on school flow and learning outcomes. The strong correlation between our proposed indicator and the IDEB highlights its reliability. Unlike the IDEB, our indicator is unaffected by state-level education policies, such as automatic promotion[3].

DATA AND METHODS

We used Grade of Membership (GoM) models, based on fuzzy clustering[4][5], to construct a multidimensional school environment quality indicator with two profiles (good vs. poor quality). We considered 37 variables related to school location, access to basic services, physical spaces, pedagogical and administrative equipment, school size, accessibility, diversity, and inclusion. The data were extracted from the annual School Censuses conducted by the National Institute for Educational Studies and Research (INEP) since 2007. GoM scores for the good quality profile range from 0 (worst) to 1 (best). Scores were calculated for each school annually and aggregated at the municipal level, weighted by student enrollment.

The IDEB, calculated every two years by INEP since 2007, combines two indicators: students' performance in standardized Portuguese and Mathematics tests under the Basic Education

---

[1] Alves, Maria Teresa Gonzaga, and Flavia Pereira Xavier. 2018. "Indicadores Multidimensionais para Avaliação da Infraestrutura Escolar: O Ensino Fundamental." *Cadernos de Pesquisa* 48, no. 169: 708-746.

[2] Allen, J. G., MacNaughton, P., Laurent, J. G. C., Flanigan, S. S., Eitland, E. S., & Spengler, J. D. (2017). Foundations for Student Success: How School Buildings Influence Student Health, Thinking, and Performance. Harvard T.H. Chan School of Public Health. https://forhealth.org/Harvard.Schools_For_Health.Foundations_for_Student_Success.pdf.

[3] Instituto de Estudos e Desenvolvimento Educacional (IEDE). 2022. *Propostas e Reflexões sobre o Novo IDEB*. https://www.portaliede.com.br/wp-content/uploads/2022/08/Propostas_Reflexoes_Novo_Ideb_Agosto2022-2.pdf.

[4] Manton, Kenneth G., Max A. Woodbury, and H. Dennis Tolley. 1994. *Statistical Applications Using Fuzzy Sets*. New York: J. Wiley.

[5] Singer, Burton H. 1989. "Grade of membership representations: concepts and problems." Pp. 317-334. In *Probability, statistics, and mathematics: papers in honor of Samuel Karlin*, edited by Samuel Karlin, et al. Boston: Academic Press.

Assessment System (SAEB) and approval rates, sourced from School Census. The index ranges from 0 to 10, with 10 being the highest performance. IDEB is calculated for both The Early Years (grades 1-5) and the Final Years (grades 6-9) of Elementary School and, more recently, for High School.

Participation in the SAEB is mandatory for public schools that meet eligibility criteria, while private schools participate through a representative sample. To ensure comparability, only public schools were considered in the calculation of the municipal average GoM score. The study analyzes the IDEB for the Early and Final Years, as well as its components – SAEB scores and approval rates. In this summary, we present only the Early Years IDEB results for brevity.

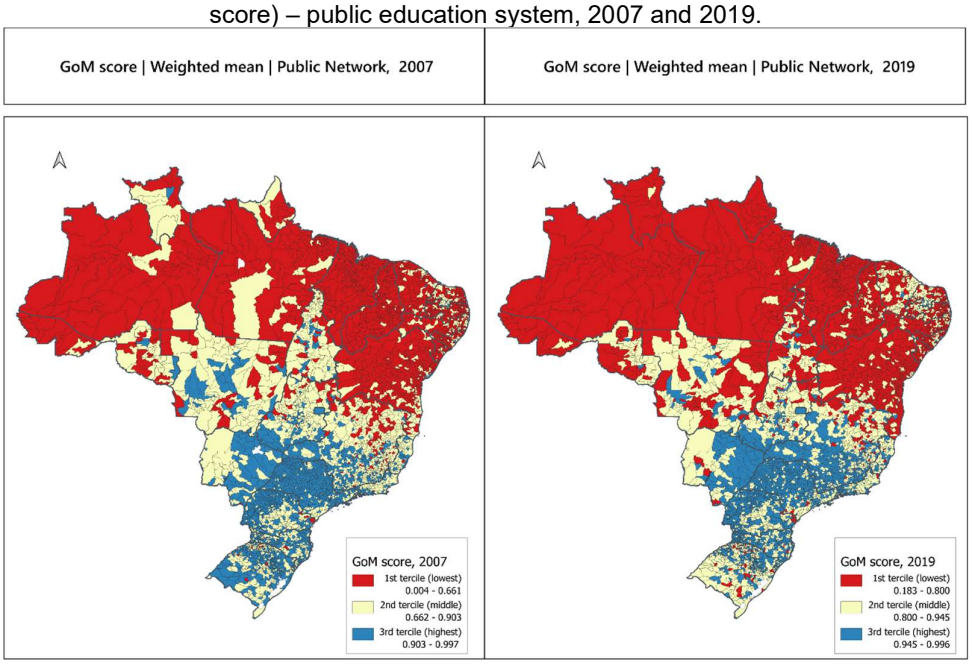
RESULTS

The maps in Figure 1 show Brazilian municipalities divided into terciles according to their GoM scores for the good quality profile, and Figure 2 shows municipalities grouped by IDEB terciles. Red represents the lowest tercile (worst case), yellow the intermediate, and blue the highest tercile (best case). Municipalities displayed in white (Figure 2) did not have their IDEB disclosed in 2019 due to insufficient participation rates (below 50%).

From 2007 to 2019, both indicators – GoM score and IDEB – improved, as evidenced by shifts in the minimum and maximum values of each tercile. In 2007, the lowest GoM scores ranged from 0.004 to 0.661, while in 2019, it shifted to 0.183 to 0.800. Similarly, the first IDEB tercile ranged from 0.9 to 3.5 in 2007, while in 2019, it increased to 2.9 to 5.3. Geographically, however, the overall distribution of quality terciles is barely changed. Municipalities in the North and Northeast continue to cluster in the lower terciles for both school environment and education quality (in red), whereas those in the South-Central regions dominate the upper terciles (in blue).

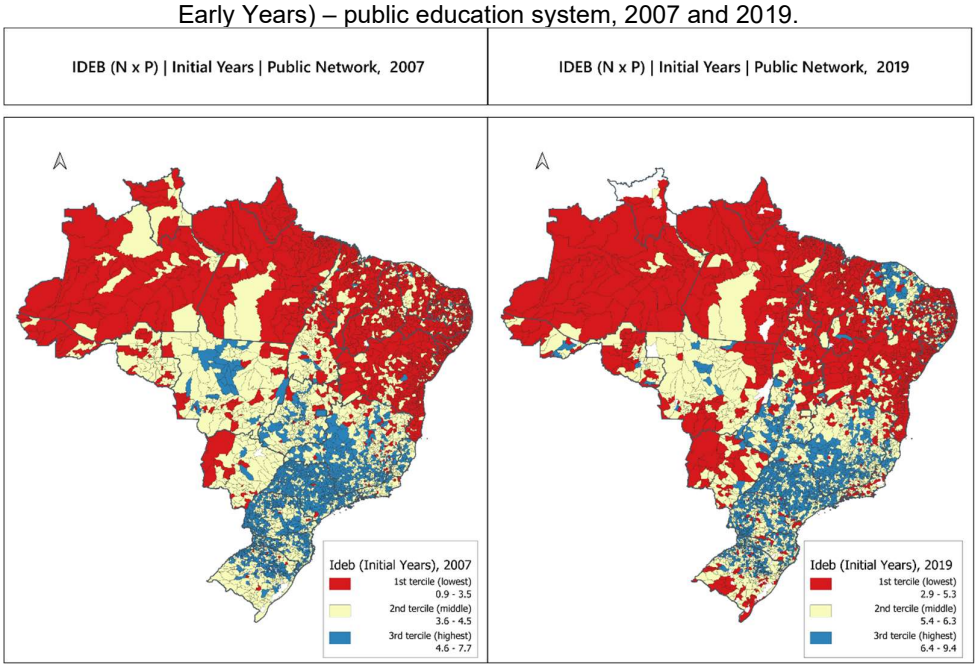
Table 1 presents descriptive statistics of school environment quality (GoM score) and education quality (IDEB) used in previous maps. A total of 5,483 municipalities were considered in 2007 and 5,419 in 2019. These tables highlight the relationship between the two variables, showing that most municipalities with low GoM score (1st tercile) also show lower educational performance, while municipalities with high GoM scores (3rd tercile) tend to perform better on the IDEB.

Figure 1 – Geographic distribution of Brazilian municipalities by terciles of School Environment Quality (GoM score) – public education system, 2007 and 2019.



Source: INEP. Brazilian School Censuses.

Figure 2 – Geographic distribution of Brazilian municipalities by terciles of Educational Performance (IDEB – Early Years) – public education system, 2007 and 2019.



Source: INEP. IDEB editions.

Between 2007 and 2019, both the school environment and education quality improved. The averages of both indicators' terciles increased, with a notable rise in the lowest terciles. The first tercile IDEB average rose from 3.02 to 4.64, a 54% increase. The average GoM score for the first tercile increased by 35%, from 0.475 to 0.643 during the same period.

Table 1 – Joint distribution of Brazilian municipalities by terciles of School Environment Quality (GoM score) and Educational Performance (IDEB) – public education system, 2007 and 2019.

Year = 2007

Variables				**IDEB [EARLY Years]**			
	Categories			1st. Tertile	2nd. Tertile	3rd. Tertile	Total
		Range		0.9 - 3.5	3.6 - 4.5	4.6 - 7.7	
			Mean	3.02	4.08	5.05	
GoM				Number of municipalities			
score	1st. Tertile	0.004 - 0.661	0.475	1,438	369	32	1,839
[Public	2nd. Tertile	0.662 - 0.903	0.800	402	892	519	1,813
System]	3rd. Tertile	0.903 - 0.997	0.957	51	652	1,128	1,831
	Total			1,891	1,913	1,679	5,483

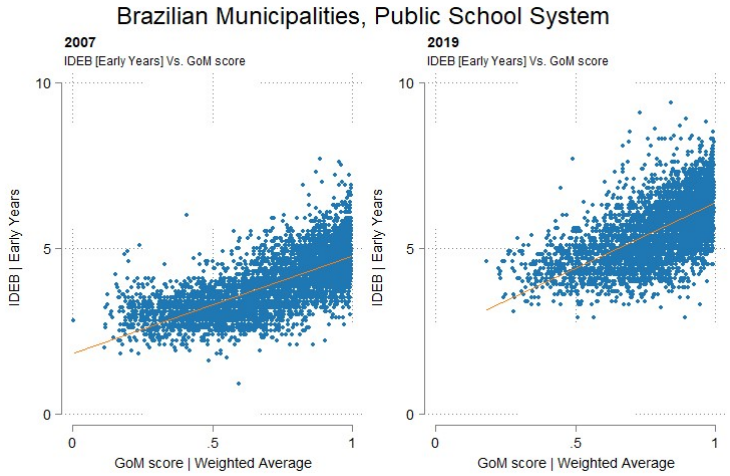
Year = 2019

Variables				**IDEB [EARLY** Years]			
	Categories			1st. Tertile	2nd. Tertile	3rd. Tertile	Total
		Range		2.9 - 5.3	5.4 - 6.3	6.4 - 9.4	
			Mean	4.64	5.89	6.82	
GoM				Number of municipalities			
score	1st. Tertile	0.183 - 0.800	0.643	1,343	370	87	1,800
[Public	2nd. Tertile	0.800 - 0.945	0.885	447	851	502	1,800
System]	3rd. Tertile	0.945 - 0.996	0.974	101	718	1,000	1,819
	Total			1,891	1,939	1,589	5,419

Source: INEP. School Censuses. IDEB editions.

Figure 3 presents scatter plots for 2007 and 2019, indicating a positive correlation between GoM score (school environment quality) and IDEB (education quality). The orange regression line reinforces this understanding. The correlation between GoM score and municipal IDEB was 0.72 in 2007 and 0.65 in 2019.

Figure 3 – Scatter plots: School Environment Quality (GoM score) and Educational Performance (IDEB) – municipal averages for the public education system, 2007 and 2019.



Source: INEP. School Censuses. IDEB editions.

EXPECTED FINDINGS

Given the evidence found so far, we expect to advance the analysis by including spatial correlation models. The aim is to verify the existence of statistically significant spatial correlation patterns between school environment quality indicators and educational performance index. The goal is to identify whether there are significant spatial distribution patterns and to map their location.

In addition to assessing spatial distribution, we also intend to explore the temporal evolution of these indicators and verify whether there is a trend towards convergence over time, particularly in the most vulnerable regions such as the North and Northeast. We expect to see a gradual convergence in school environment quality and educational performance, although regional disparities may persist for an extended period.

The results of the spatial correlation models will be useful in identifying clusters of municipalities with similar characteristics, providing valuable insights for more targeted public policies to address regional inequalities.