Public Policies and Learning Inequalities by Socioeconomic Level: How Much Can We Level the Playing Field?

Samuel Franco[1] **and Flavio Riva**[2]

[1]Oppen Social.
[2]Institute for Mobility and Social Development (Imds).

Abstract This technical note presents an analysis of the socioeconomic differences in Portuguese language learning among students from public schools in Brazil. We used two data sources for the period from 2009 to 2023: (i) microdata from the Basic Education Evaluation System (SAEB) for the 5th year of elementary school; (ii) data from the socioeconomic index of the schools where these students studied when they took the test. The combination of this information allows us to consider whether and to what extent the period was marked by a reduction in socioeconomic inequalities.

The note documents significant differences in the rate of students who achieved adequate performance at different points in the distribution of the socioeconomic index. For example, in 2023, only 3 out of every 10 students in the bottom tenth of the index reached this level, compared to 8 out of 10 in the top tenth. Between 2009 and 2023, all deciles increased by a similar amount, averaging 1.6-1.9 percentage points per year. As a result, the socioeconomic differences in learning are, in the most recent data, similar to those of 2009: the deciles have shifted, but these changes have not been accompanied by a parallel reduction in inequalities.

As Ceará is considered an exemplary case in literacy and in elementary education, we restricted the analysis to students from public schools in that state and characterized a process of convergence between the tenths of the socioeconomic index, concentrated temporally between 2009 and 2015 —the period in which students, affected by an innovative set of state policies to improve education, reached the year of sitting SAEB exams. Between 2009 and 2015, socioeconomic inequalities in learning deepened among students in public schools in other Northeastern states, in a manner similar to that observed in Brazil as a whole.

1 Introduction

The main objective of this technical note is to analyze the extent to which Brazil has been able to reduce educational inequalities as it increases learning levels. It is true that the country has historically faced challanges in improving educational performance at the same pace as it increases the number of full years of study. However, as we will see, performance improved considerably relative to that achieved at the end of the 2000s. The question that this note proposes to address is the extent to which these advances have been accompanied by reductions in learning inequalities.

To analyze the issue, we focused our attention on the first stages of schooling and on the Portuguese Language countrywide standardized exams. The justification for these choices is due to (i) the fact that inequalities at the beginning of the learning cycle can be perpetuated, or even expanded, during subsequent cycles, and (ii) the importance of reading and writing skills for learning in other subjects, especially when treated at the stage in which literacy is consolidated.[1]

[1]Ideally, the analysis would begin with the results obtained at the end of this cycle, in the 2nd year of elementary school. However, there are no data series long enough to enable long-term national-level analysis with sufficient precision. The first school year that allows a descriptive analysis

The note has three sections, in addition to this introduction. Section 2 discusses the databases used and provides contextual information to support the interpretation of the main results. Section 3 presents these results. Section 4 concludes the note.

2 Data and Context

2.1 Achievement in Portuguese Language on the SAEB

In the 5th year, proficiency in Portuguese Language in the SAEB is evaluated on a scale of 100 to 400. Cuts on this scale define discrete jumps in the probability that the student has reached fundamental skills in the test's reference matrix (Bof, 2022). In this note, we used the adequate level, defined as proficiency greater than or equal to 200, as the primary variable of interest. Proficiencies in this interval indicate that students can, for example, infer cause-and-effect relationships and differentiate facts from opinions in informative and narrative texts. Reaching the adequate level, therefore, constitutes an important signal of the consolidation of skills that characterize interpretative reading, a determining condition for learning and for the student's future school trajectory.

Figure 1 shows the percentage of students from public schools that reached this milestone for Brazil, the Northeast (excluding Ceará), and Ceará, between 2007 and 2023. Between 2007 and 2017, Brazil increased the proportion of children with adequate learning in public schools from 27% to 55%. Ceará, in turn, showed an average annual improvement of 5 percentage points, starting from a level very close to that of the Northeast and surpassing Brazil between 2013 and 2015.

It is possible to trace the origins of the gains in Ceará to the implementation of public policies aimed at early elementary education. The first of these initiatives was the Literacy at the Right Age Program (PAIC), implemented as a state public policy in 2007. With the objective of ensuring that all children in public schools were literate by age 7, PAIC implemented systemic changes in the municipal and state bureaucracies responsible for literacy, beginning in 2007, including new forms of monitoring the development of essential skills and continuing education programs for literacy teachers.

In addition, the prioritization of literacy at the right age by the municipalities was stimulated by quotas of the Tax on the Circulation of Goods and Services (ICMS) to municipalities for education, health, and environmental indicators, with the greatest weight for indicators related to literacy (de Gusmão e Ribeiro, 2011). In subsequent years, other programs extended the scope and/or enhanced PAIC. This is the case of the 2009 Nota Dez School Program, which established mentoring between principals of low- and high-performance schools and financial incentives, and the 2011 PAIC +5, which extended PAIC's target audience to students in the 3rd, 4th, and 5th grades of elementary school.

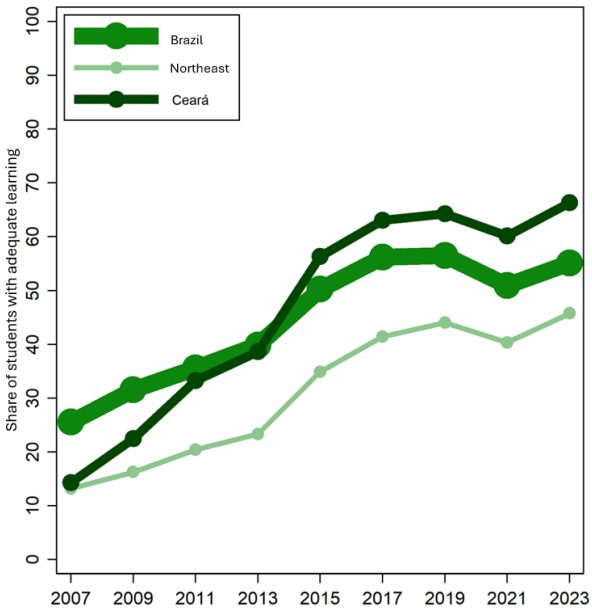
Note that the students who took the SAEB tests in 2013 and 2015 entered elementary education in the period of expansion and consolidation of this set of initiatives. This variation in exposure over time allowed the literature under evaluation to document evidence of the impact of these public policies (for example, in Costa & Carnoy, 2015, on Paic, Goldemberg et al., 2021, on Prêmio Escola Nota 10, e Lautharte et al., 2021, on the distribution of the ICMS, conditional on results). This note draws on the literature to propose a characterization of the evolution of students' performance under these initiatives, incorporating information on their position in the socioeconomic distribution.

2.2 Socioeconomic status

The second source of information used is the socioeconomic index of schools and students. To create the historical series between 2009 and 2023, we followed the steps described below: The socioeconomic index of the schools was made available by INEP from 2011 to 2021. The measure is calculated from the simple arithmetic average of the socioeconomic index of students enrolled in a given school. The stu-

of this nature is the end of elementary school (5th year), the focus of this study.

Figure 1. Share of Students with Adequate Portuguese Language Learning in the 5th year of Elementary School, Brazil, Ceará and Northeast (2007 to 2023)



Nota: The figure describes the percentage of public school students who reached an adequate level in Portuguese language per year of the SAEB test, for Brazil, for the Northeast (without Ceará) and for Ceará, from 2007 to 2023.

dent index is calculated based on the students' answers to the contextual questions of the SAEB and the National High School Exam (Enem) assessments and which only became available in microdata as of 2023. The questions used concern family income, ownership of assets and the hiring of services of domestic employees by the students' families and the level of education of their parents or guardians.[2]

Using the schools' INEP codes in the students' microdata, it is possible to combine the performance information with information on the socioeconomic index of their schools, from 2011 to 2021. There are no public data for the years prior to 2011. For this reason, we assigned to the students of the 5th grade of 2009 values of the socioeconomic index of the schools in 2011.[3] At the time of this note's release, the 2023 school index data had not yet been made publicly available on the INEP website. For this reason, in 2023 we used individual student data, which, as discussed above, was publicly released for the first time this year. More specifically, using the microdata, we computed simple arithmetic means of the student index with the same identifier as that of the school.

Using the school identifiers, the information of the schools' socioeconomic index was connected to the microdata from 2009 to 2023, and deciles of the variable were constructed for Brazil, the Northeast (without Ceará), and Ceará.[4]

3 Analysis

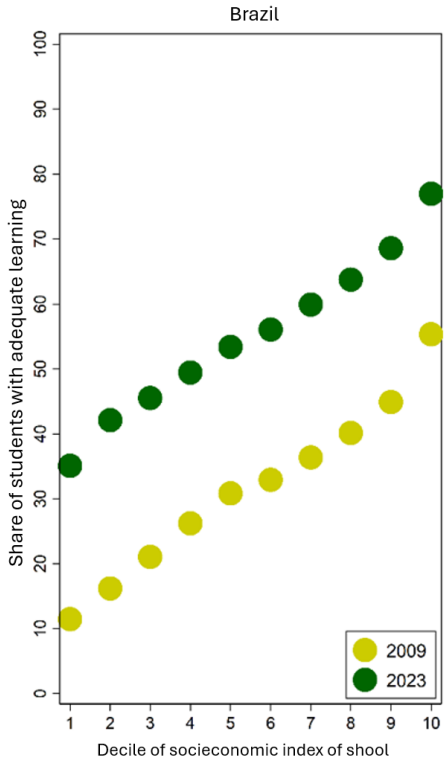
The main goal of this note is to characterize the evolution of Portuguese language learning in the early years of elementary school, incorporating into the analysis considerations about the in-

[2]The methodology to calculate the index uses Item Response Theory to construct the indicator, synthesizing the information in a single numerical score for each student (INEP, 2011). The index is expressed in a continuous scale, with a mean of 50 and a standard deviation of 10, where higher values indicate a higher socioeconomic level. There are differences in the variables used by INEP over the years to improve the indicator (INEP, 2013, 2015).

[3]The combination of the data generates a loss of less than 10% of the students who took the test, for all years and geographic units. The procedure was not repeated in 2007 because there was a significant loss (more than 30%) in the database integration process.

[4]Dispersion measures were computed using the empirical distribution of the school's socioeconomic index in the microdata, restricted to a given area in a given year. Thus, the correct interpretation of a decile always takes the geographical unit as a set of reference; for example, the 1st tenth in Ceará, in 2009, represents the group that had socioeconomic index values among the lowest 10% in the group of students from public schools in Ceará in that same year.

Figure 2. Adequate learning of Portuguese Language in the 5th grade, by tenths of the socio-economic index (public schools), 2009 and 2023



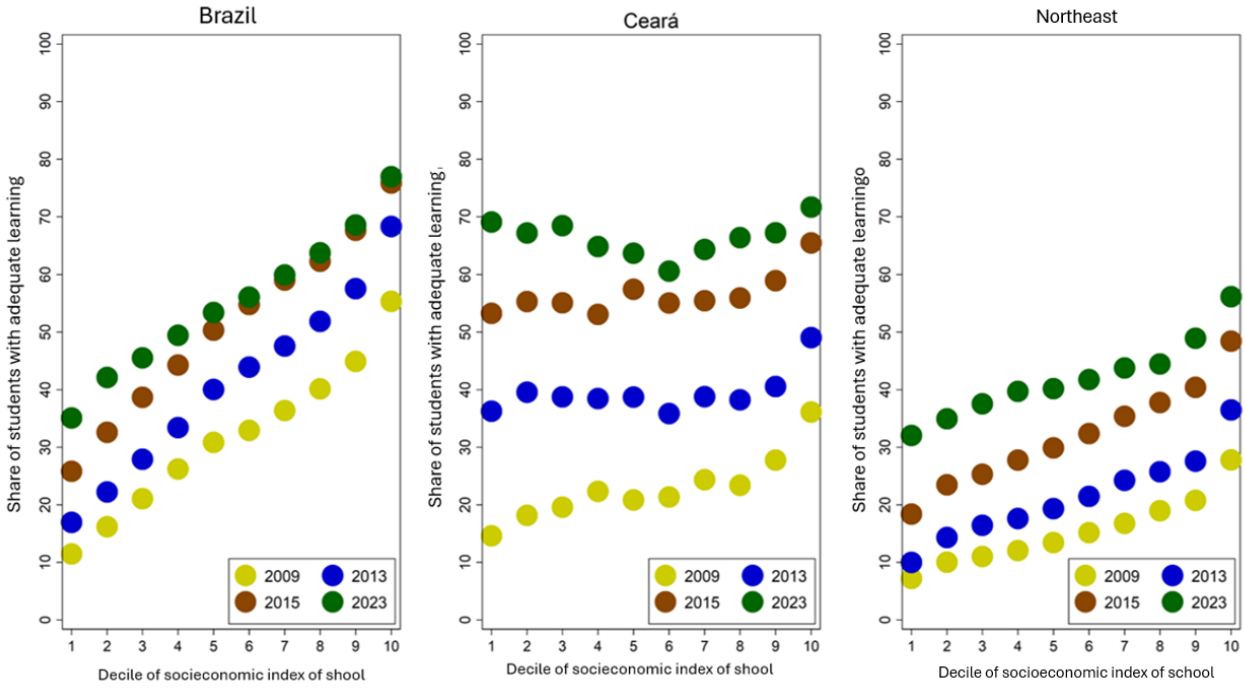
Note: The figure above was constructed using microdata from the Basic Education Evaluation System (SAEB) for the 5th year of elementary school; (ii) data on the socioeconomic index of these students. Each point on a graph represents the percentage of students who had adequate performance in Portuguese for a given tenth of the socioeconomic index.

equality of performance of students from different socioeconomic profiles. To do so, we used the base described above to compute the percentage of students with an adequate level by socioeconomic index decile.

Figure 2 shows a comparison of Brazil in 2009 and 2023. In both years, there are significant differences in the percentage of students who reached the appropriate level of learning in each tenth of the index. For example, in 2023, only 35.1% of the bottom tenth reached this level, compared to 77.1% in the top tenth; in 2009, the figures were 11.5% and 55.3%, respectively. A difference in level similar to that found for the extreme tenths separates the intermediate tenths between the two periods. Columns (5) and (6) of Panel A, in Table 4, show that this difference ranges from 21.6 to 25.9 percentage points, and that all tenths grew at average rates of 1.6 to 1.9 percentage points per year. This suggests that the country's public schools reveal, in the most recent data, a similar capacity to level the playing field of reading at the end of the early years for students in schools with different socioeconomic profiles.

Figures 3 and 4 allow us to establish comparisons between the long-term results found for Brazil and those found for students in public schools in Ceará and the Northeast. In addition, we also present medium-term results for 2013 and 2015, years in which children affected by the public policies described in Section 2 reached 5th grade. Note that, in 2023, Ceará managed to practically eliminate performance differences between socioeconomic levels. The process of approximation between the tenths occurred because, between 2009 and 2023, the state significantly improved students' performance from lower socioeconomic backgrounds. In the period, the average annual growth of the first five tenths was 3.1 to 3.9 percentage points, compared to 1.6 to 1.9 for Brazil as a whole, and 1.7 to 1.9 for the Northeast (see column column (6) in Tabelo 4). Finally, Ceará was able to raise the performance of its students of lower socioeconomic status much more significantly compared

Figure 3. Adequate learning of the Portuguese language in the 5th grade, by tenths of the socio-economic index of the schools



Nota: The figures above were constructed using two sources of information for the period from 2009 to 2023: (i) microdata from the Basic Education Evaluation System (SAEB) for the 5th year of elementary school; (ii) data on the socioeconomic index of the schools where these students studied when they took the test. Each point in a graph represents the percentage of students who had adequate performance in Portuguese for a given tenth of the socioeconomic index, in the geographic unit highlighted at the top of the figure. The exercise for the Northeast was carried out without including students from schools in Ceará. Section 2 presents more details about the databases and the indicators used.

to those of higher level between 2009 and 2013 and between 2009 and 2015 (see columns (7) and (8) in 4).

For joint visualization of movements in learning levels and in the distribution by tenths, Figure 5 represents the means and standard deviations between tenths for each geographic unit and year, between 2009 and 2023 (excluding 2021). Both Brazil and the Northeast have a similar pattern of evolution, with an inverted-shaped U curve. In these geographic units, inequality between tenths increases by a similar magnitude between 2009 and 2015, remains practically stable between 2015 and 2019, and then falls in 2023. For Ceará, on the other hand, there is a downward movement in inequality and a greater increase in averages than that found for Brazil and the Northeast of Brazil.

4 Final Comments

This technical note examines the evolution of socioeconomic levels and inequalities in Portuguese language learning at the elementary school level in public schools in Brazil between 2009 and 2023. We document that the inequality of students of higher and lower socioeconomic status in Brazil is, in the most recent data, similar to that of 2009. On the other hand, we show that there has been, in Ceará, a process of convergence between the tenths of the socioeconomic index and a substantial increase in its level, especially between 2009 and 2015 – a period in which students affected by an innovative set of state policies to improve education reached the year of the SAEB tests.

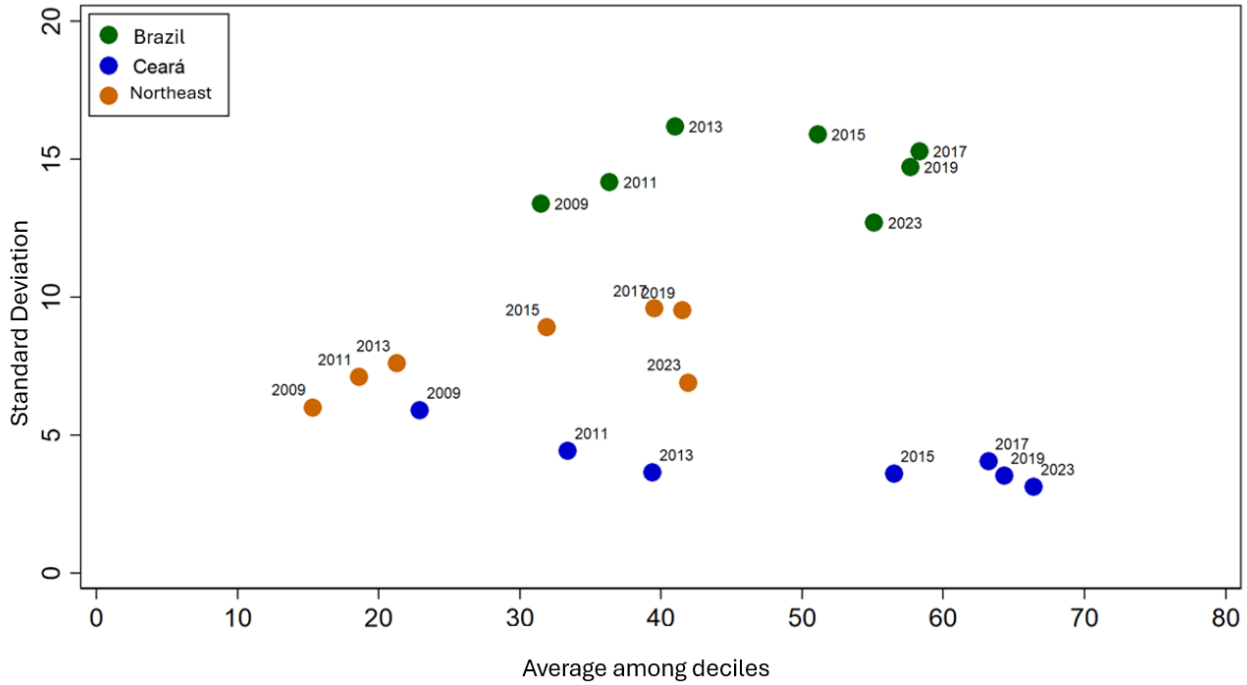
There are other topics that were not addressed by this note and that may be the object of future research efforts. First, literacy on the PAIC and other public policies of the period documents average effects, but there is less atten-

Figure 4. Levels and growth of adequate learning of the Portuguese language, by tenths of the socioeconomic index of schools

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Variation (in p.p.s)	Average annual variation (in p.p.s)		
	2009	2013	2015	2023	2009-2023	2009-2023	2009-2015	2009-2013
	Panel A. Brazil							
1	11,5	17,0	25,8	35,1	23,6	1,69	2,39	1,38
2	16,2	22,2	32,6	42,1	25,9	1,85	2,74	1,51
3	21,1	27,9	38,7	45,5	24,4	1,75	2,93	1,71
4	26,2	33,4	44,3	49,4	23,2	1,66	3,01	1,80
5	30,9	40,0	50,3	53,4	22,6	1,61	3,25	2,29
6	32,9	43,9	54,8	56,1	23,2	1,65	3,64	2,75
7	36,4	47,6	59,1	59,9	23,5	1,68	3,78	2,79
8	40,1	51,9	62,3	63,8	23,6	1,69	3,69	2,93
9	44,9	57,5	67,7	68,6	23,7	1,69	3,79	3,16
10	55,3	68,3	75,9	77,0	21,6	1,55	3,42	3,25
	Panel B. Ceará							
1	14,6	36,3	53,2	69,1	54,5	3,89	6,44	5,41
2	18,2	39,6	55,3	67,2	49,1	3,50	6,19	5,35
3	19,6	38,8	55,1	68,5	48,9	3,49	5,91	4,79
4	22,3	38,4	53,1	64,8	42,5	3,04	5,12	4,03
5	20,9	38,7	57,4	63,7	42,8	3,06	6,10	4,46
6	21,4	35,9	55,0	60,6	39,2	2,80	5,61	3,63
7	24,4	38,8	55,4	64,4	40,0	2,86	5,18	3,61
8	23,4	38,2	55,9	66,4	43,0	3,07	5,42	3,70
9	27,8	40,6	58,9	67,2	39,5	2,82	5,20	3,20
10	36,1	49,0	65,5	71,7	35,6	2,54	4,90	3,24
	Panel C. Northeast							
1	7,2	10,0	18,4	32,0	24,8	1,77	1,86	0,69
2	10,0	14,3	23,5	34,9	24,9	1,78	2,24	1,07
3	11,0	16,4	25,3	37,5	26,5	1,89	2,38	1,35
4	12,1	17,6	27,8	39,7	27,6	1,97	2,61	1,39
5	13,4	19,4	29,9	40,2	26,8	1,91	2,74	1,48
6	15,2	21,5	32,4	41,8	26,6	1,90	2,87	1,57
7	16,8	24,3	35,4	43,8	27,0	1,93	3,10	1,87
8	19,0	25,8	37,8	44,5	25,5	1,82	3,13	1,69
9	20,8	27,6	40,4	48,9	28,1	2,01	3,27	1,70
10	27,8	36,5	48,4	56,1	28,3	2,02	3,44	2,17

Nota: The table above was constructed using two sources of information, for the period from 2009 to 2023: (i) microdata from the Basic Education Evaluation System (SAEB) for the 5th year of elementary school; (ii) data on the socioeconomic index of the schools where these students studied when they took the test. Columns (1) to (4) present the percentage of 5th grade students who had adequate performance in Portuguese in the geographic units indicated in the whole of each panel, grouping students in tenths of the average socioeconomic index of their schools. Columns (5) to (8) present summary measures of the variation in the indicators between the years

Figure 5. Evolution of the mean and standard deviation between tenths over time, by geographic unit



Notes: The figure above represents the joint movement of the mean of the deciles and the standard deviation of each geographic unit over time.

tion to the effects on the distribution of learning and, in particular, effects on socioeconomic inequalities. Reviews that might contemplate this dimension can be important for a more comprehensive consideration of the importance of initiatives aimed at basic education and, in particular, literacy. Second, the results do not consider changes in composition resulting from fluctuations in dropout, abandonment, and failure rates. A more precise investigation of the evolution of engagement with the school system using the excerpts from this note can help to clarify the weight of these considerations for the interpretation of the results, in addition to providing important information about other indicators of educational inequality. Finally, it would be important to investigate to what extent inequalities in the learning of the Portuguese language during junior high school are associated with levels of content learning and inequalities in other subjects in the same cycle (e.g., in Mathematics) and in subsequent cycles.

Suggested Citation

Franco, S., Riva, F. (2026). *Public Policies and Learning Inequalities by Socioeconomic Level: How Much Can We Level the Playing Field?*. Technical Note ImdsNT004-2025. Rio de Janeiro: Institute for Mobility and Social Development

References

Costa, L. O., & Carnoy, M. (2015). The effectiveness of an early-grade literacy intervention on the cognitive achievement of Brazilian students. *Educational Evaluation and Policy Analysis*, *37*(4), 567–590.

Goldemberg, D., Bacalhau, P., & Junior, I. J. L. (2021). Parcerias com incentivos podem melhorar escolas de baixo desempenho? Evidências do estado do Ceará.

INEP. (2011). Nota Técnica-Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse).

— (2013). Nota Técnica-Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse).

INEP. (2015). Nota Técnica-Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse).

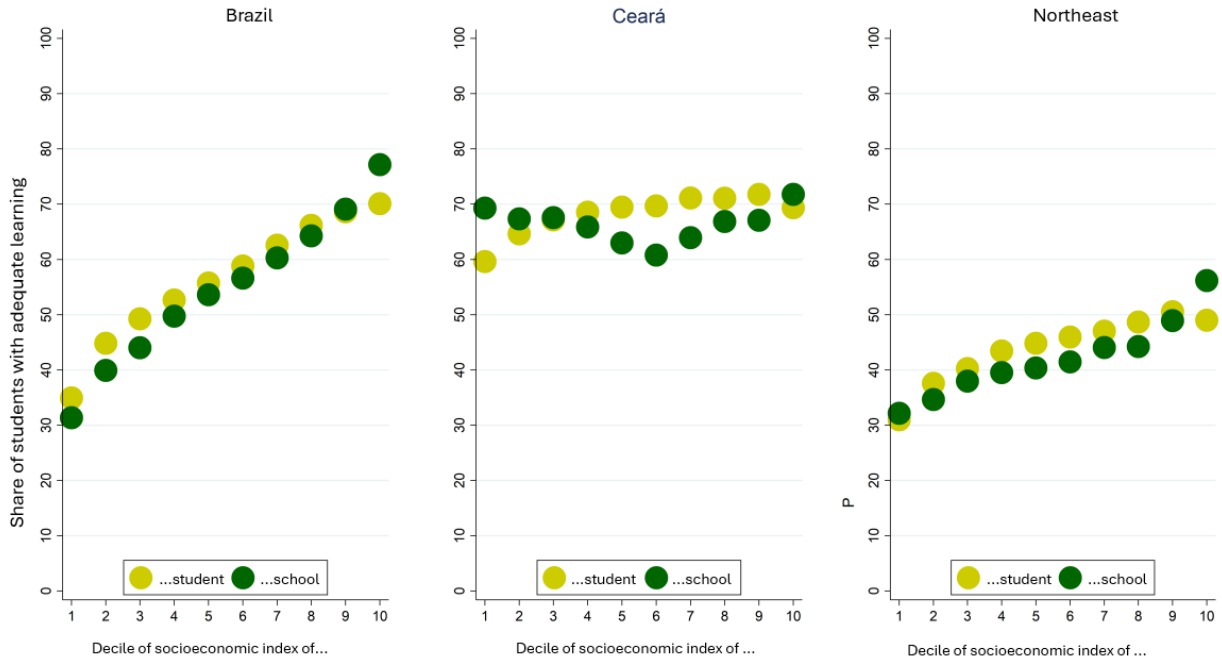
Lautharte, I., De Oliveira, V. H., & Loureiro, A. (2021). Incentives for Mayors to Improve Learning. *World Bank, Washington, DC*.

Appendix: Achievement Gradient in 2023

The main results of the note use the socioeconomic index of schools, publicly available by INEP since 2011. However, it is important to recognize that, when we identify a student with the same average as that of other students in the school, with the aggregate data made available by INEP, we lose important information about the environment of material and non-material goods of the students' families. As discussed in the text, in 2023, in an unprecedented way, the SAEB data included information on the socioeconomic index of the student's family.

The purpose of this appendix is to use individual data to compute the tenths and compare them with the results presented in the note for 2023. .6 shows the results of the exercise. It is interesting to observe that both curves have a similar slope in the national data and clash strongly only in the upper tenth. The figure also shows that, although there is still a positive slope between performance and individual socioeconomic index in the lower tail for Ceará, it is substantially lower than that of Brazil and the Northeast. In addition, in Ceará, students from public schools of all tenths have rates equal to or greater than 60%, which is achieved only by students in the 6th decile in Brazil and by no decile of students in the Northeast.

Figure .6. Adequate learning of Portuguese Language in the 5th grade, by tenths of the socio-economic index (public schools), 2009 and 2023



Nota: The figures above show the percentage of students who performed adequately in Portuguese for each decile of the student and school index, using microdata from the 2023 SAEB. The exercise for the Northeast was carried out without including students from schools in Ceará.