WORLD BANK GROUP
Education

Can peer mentoring coupled with incentives affect school turnaround? Evidence from Ceará state in Brazil

Diana Goldemberg

Priscilla Bacalhau

Ildo José Lautharte Junior

June 30, 2020

Abstract

This note presents a peer-mentoring program for principals in Brazil, *Prêmio Escola Nota 10,* and discusses its impacts on second-graders learning outcomes. The program matches the principals of schools with poor academic performance with the principals from top performance schools and both receive a small grant. As an incentive for active participation, a second disbursement of the grant is provided to both schools, conditional on the subsequent improvement of the low-performing school in standardized exams. Exploiting this design, regression discontinuity regressions using school ranks indicate that low-performing schools participating in the program increase 0.18 standard deviations the education quality indexes in the following year. The effects for low-performing participants are positively associated with school size and geographical proximity between paired schools. Participation does not significantly impact top-performing schools. These findings illustrate how school turnaround may be achieved systemically.

Keywords: education performance, school turnaround, principal peer-mentoring

1. Introduction

Turning around low-performing schools is a perennial challenge for education systems in developed and developing countries alike. Despite numerous well-documented cases of successful school turnaround, there is a knowledge gap on systemic and scalable approaches to school improvement[1]. School principals are often considered the main lever for school turnaround[2], but most professional development and mentoring programs for school leaders are restricted to newly appointed or aspiring principals[3]. Principals from the worst-performing schools are often mandatorily replaced in turnaround policies and programs[4].

Ceará's experience in systematically turning around low-performing schools stands out, even without principals' replacement. In this relatively poor state in Northeast Brazil[5], educational outcomes have dramatically improved in recent years. Only 7% of Ceará's municipalities scored above Brazil's average in the national quality index for primary education[6] in 2007, soaring to 71% a decade later. While this success cannot be attributed to any single policy or program, it is grounded on a series of reforms towards results-based financing and the provision of technical assistance to the municipalities' education boards[7]. The backbone of these initiatives was the Literacy Program at the Right Age (*Pacto pela Alfabetização na Idade Certa*, PAIC), implemented in 2007[8]. Though the program focuses on early-grade literacy, it played a major role in improving learning outcomes at the end of primary education[9]. It includes the distribution of teaching materials, face-to-face training to early-grade teachers, a state-wide annual standardized assessment, and a peer-mentoring program for principals to award the top-performing schools while supporting the low performers.

This note describes Ceará's incentivized peer-mentoring initiative, *Prêmio Escola Nota 10* (PEN10), and some evidence of its effectiveness. We focus on this single component of PAIC, described in detail in the next section. PEN10 represented more than half of PAIC's budget in 2017[10] and its design allows for a robust impact assessment. Mentoring programs for school principals are not a new idea[11], but impact

[1] Herman (2012); Baroody (2011); Herman et al. (2008); Sternberg et al. (2006)

[2] Meyers and Sadler (2018), Wikeley et al. (2005)

[3] Spiro, Mattis, and Mitgang (2007); Service, Dalgic, and Thornton (2018)

[4] Such as the federally funded School Improvement Grants in the US (Trujillo and Renée, 2015).

[5] In 2017, monthly per capita income in Ceará was R$ 824, roughly US$ 200, and 65% of the Brazilian national average (IBGE, Pnad 2017).

[6] Index of Basic Education Quality (*Índice de Desenvolvimento da Educação Básica*, IDEB) of elementary municipal schools, composed by the performance of students in national assessments at the end of fifth grade with progression rates.

[7] Loureiro et al., 2020

[8] PAIC was a scale-up of a successful early-grade literacy initiative led by Sobral municipality in 2001, by the mayor that later became Ceará's governor and created the state-wide program. In 2012, the Brazilian Federal government attempted to replicate it at a national scale, under the name PNAIC.

[9] Costa and Carnoy (2015) estimate a causal impact from PAIC in Portuguese and Math scores (0.10 and 0.18 standard deviation, respectively) in grade 5, using triple differences.

[10] PEN10 cost was R$ 43 million in 2017, representing 52% of PAIC's budget. PAIC was complemented by fiscal incentives to municipalities, tied to their educational performance, which are not included in the program's budget, and represented roughly R$ 200 million in 2009 and R$ 513 million in 2017.

[11] They were the most cited priority in a 2017 survey of U.S. state education board leaders (Riley and Meredith, 2017).

evaluations of such programs are rare, and evidence is mixed[12]**.** This work builds on Goldemberg (2019), which uses a regression discontinuity design to estimate significant learning gains in low-performing schools due to program participation. The next sections describe the program, methodology and main results, as well as suggestions for future analysis. Portraying and assessing this innovative program are key steps for understanding its benefits and possible expansion to other territories or contexts, considerations raised in the last section.

2. Ceará's peer-mentoring initiative for principals: *Prêmio Escola Nota 10*

The peer-mentoring program implemented in Ceará pairs school principals with poor academic performance with principals from top-performing schools and provides a small grant to both. PEN10 is the evolution of school incentive programs that started in 2001[13]. PEN10's execution relies on a well-established learning assessment system (*Sistema Permanente de Avaliação da Educação Básica do Ceará*, SPAECE) that is census-based, annual and transparent. The state publishes disaggregated results for all public schools and municipalities, and a consolidated Education Quality Index (EQI) is the main metric for program participation. Since 2009, the PEN10 program selects[14] the 150 top and bottom schools from the EQI ranking within a target grade, provided they meet certain eligibility criteria. Pairs of top- and low-performing schools are matched[15], and all participants receive an unconditional grant when they join the program. Instead of promoting principal replacement, the program provides resources enabling mentorship between the paired schools' current leadership teams. Over the following year, the principal from the top-performing school should provide mentoring for the paired low-performing school, aiming to improve its subsequent results. The teams are expected to meet at least six times. In the next assessment, if the low-performing school successfully improved its EQI, among other conditions, both schools receive a second disbursement of the grant[16]. Figure 1 details the implementation steps of PEN10 in Ceará. Initially, the program focused on the assessment of second graders, but currently, fifth and ninth graders learning outcomes define participation. This note focuses on the 2009 to 2015 editions of PEN10 for which participation was based on the second graders' literacy assessment (SPAECE Alfa)[17].

---

[12] Herman et al (2017) present a comprehensive review of school leadership interventions under the *Every Student Succeeds Act* and find that only two US support programs have been subject to experimental or quasi-experimental impact evaluation studies. The conclusions of those two studies are mixed. The *McREL's Balanced Leadership Program* was found to increase staff stability in treated schools but had no effect on student achievement (Jacob et al., 2015). The *National Institute for School Leadership Executive Development Program* showed positive effects on reading and math achievement (Nunnery et al., 2011).

[13] The predecessors *Escola do Novo Milênio* (New Millennium School Award, active from 2001 to 2004) and *Prêmio Escola Destaque* (Highlight School Award, active from 2004 to 2009, replaced by PEN10) also provided monetary awards to top-performing schools. However, neither offered monetary grants nor peer-mentoring to low-performers, core to PEN10.

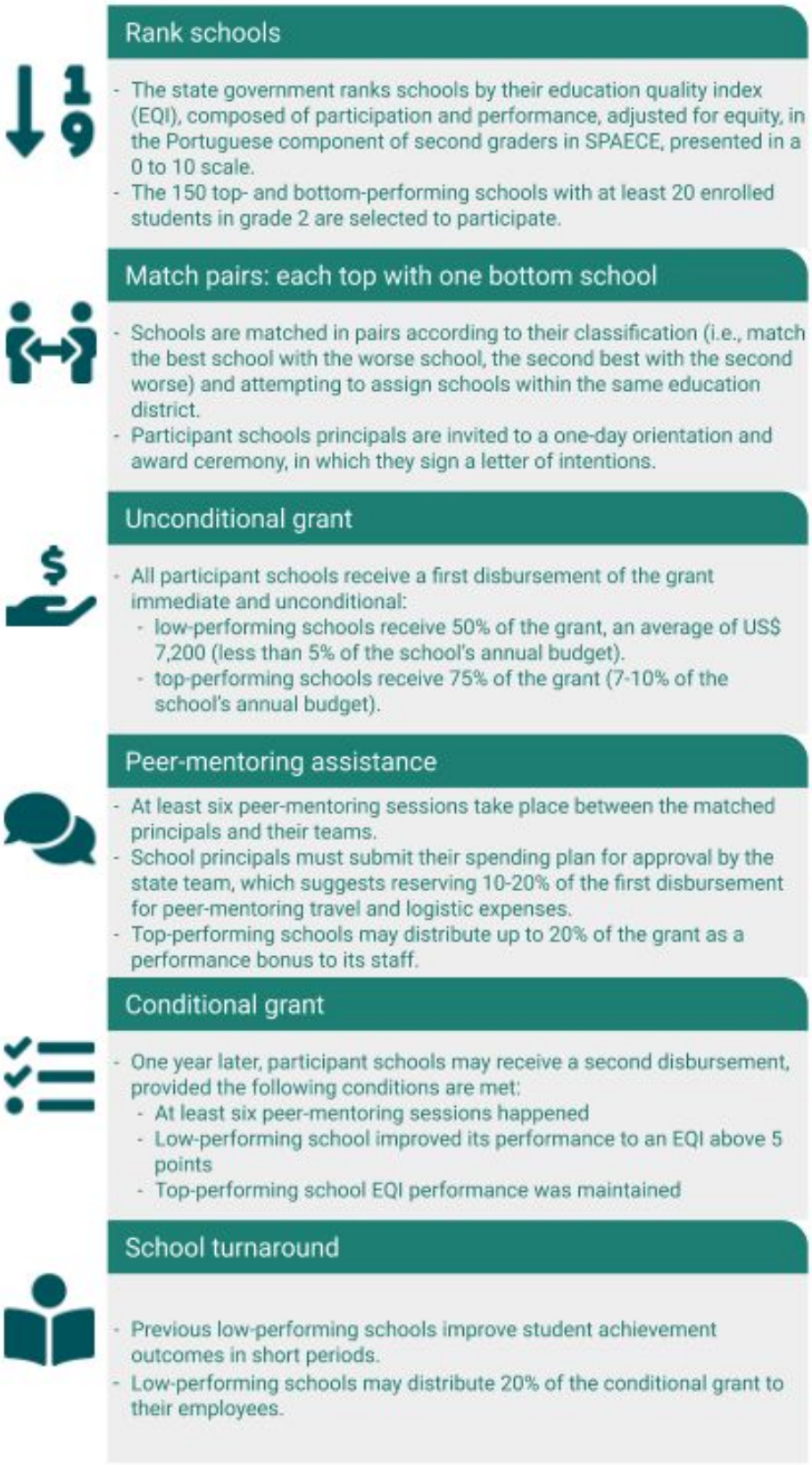
[14] Though selected principals could decide not to participate, the take-up is 100% to date.

[15] There was not a formal algorithm to match the schools, but the state education board reports using both the classification (i.e., match the best school with the worst school, the second best with the second worst, etc.) and attempt to assign schools within the same education district.

[16] For the low-performing schools, each disbursement of the grant is, in average, US$ 7,200 (less than 5% of the school's annual budget). The top-performing schools receive 75% of the award in the first disbursement and the remaining 25% in the second one, conditional on having successfully helped their mentee to improve, while sustaining their own high performance.

[17] Despite discontinuing the mentorship component based on 2nd-grade results and selection of low-performing schools, the program continues to award the 150 top performers in this grade-level.

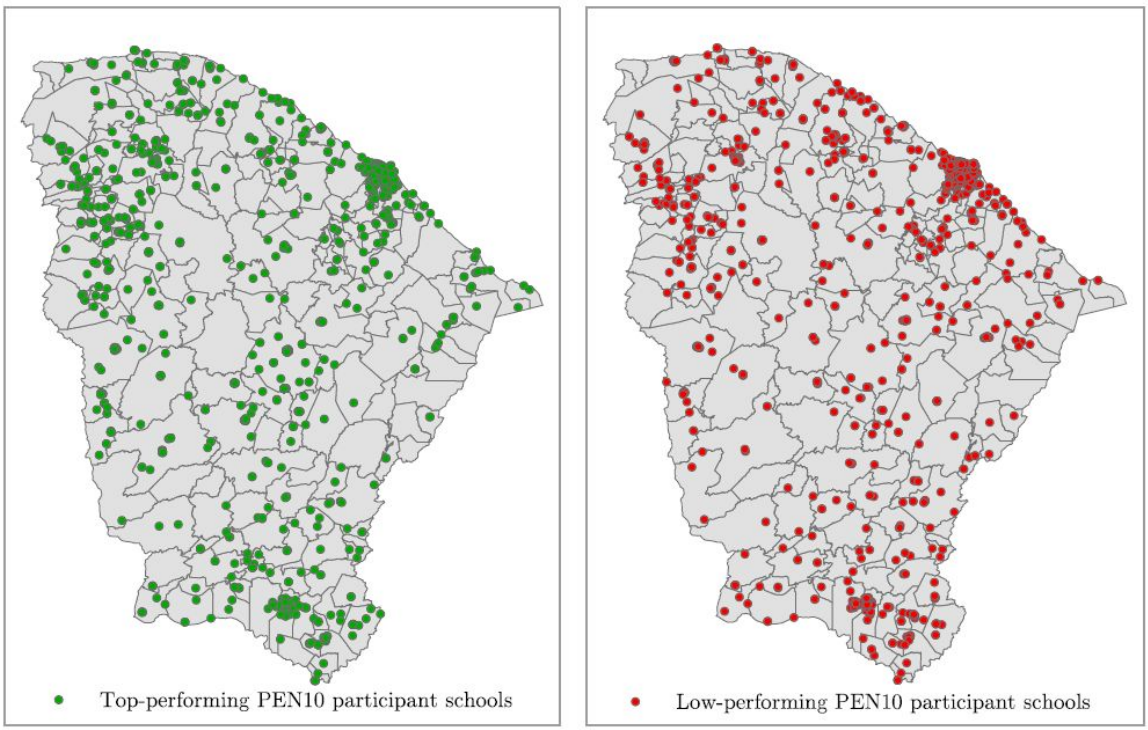
Figure 1 - How the PEN10 works



Source: Authors' elaboration.

Participant schools were spread out throughout the state territory, both for top and bottom schools. The maps in Figure 2 show the geographic distribution of the top-performing and low-performing PEN10 participant schools in the Ceará, considering all participants from 2009 to 2015. The maps indicate that top- and low-performing schools may be found within the same municipality, although this was not required by the program design. The lined-up geographic distributions may facilitate a local matching of the pair of schools.

Figure 2 - Map of participant schools

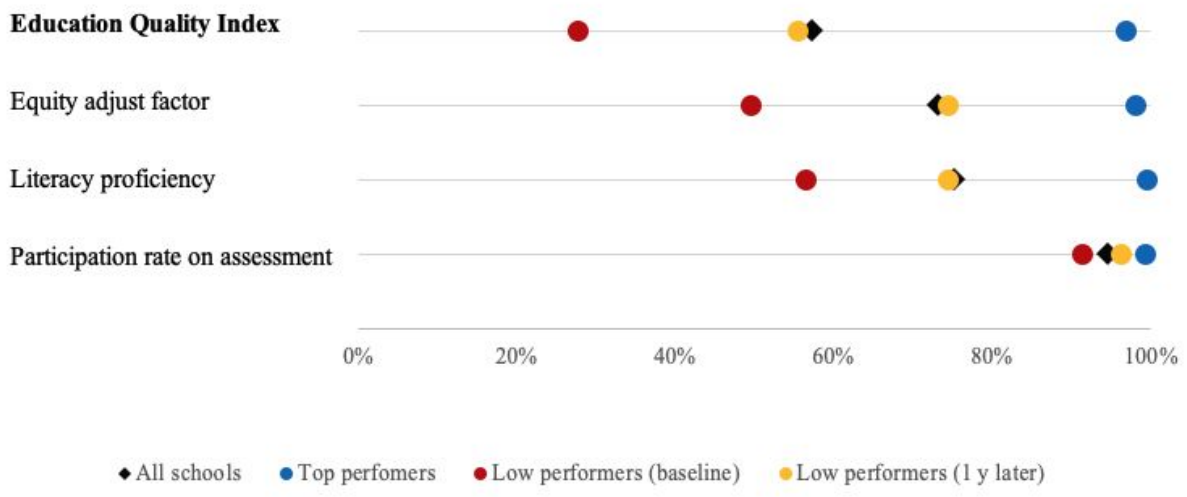


Source: Authors' elaboration based on Ceará Education Board (SEDUC)

By design, top-performing schools had better learning outcomes than the low-performers, but they also differed in other aspects. In terms of performance, the baseline Education Quality Index (EQI) of top-performing participants (9.7 in a 0 to 10 scale) is higher than observed for the low-performing participants (2.8). After one year of participating in the peer-mentoring sessions, the low-performing schools moved towards the state average EQI. Despite the general improvement of Ceará's performance in the recent years, the empirical design vanishes out the influence of such trend from the estimates of PEN10 impacts. This catch-up happens in all three components of the EQI, namely, average literacy proficiency, equity factor, and participation (see Figure 3). Top- and low-performing schools differ in other significant ways, that are not expected to change due to the participation in the program: the top performers are smaller (39 versus 52 students enrolled in the second grade), have a larger share of students using free school transport (28% versus 18%) and have a lower share of students below their ideal grade (9% versus 23%). Surprisingly, teacher qualification, measured by the share of teachers with a

bachelor's degree, and school infrastructure[18] are significantly worse in the top-performing schools. In Appendix 2, Table 1 presents descriptive statistics of the participant schools.

Figure 3 - Performance in the State Assessment



Source: Authors' calculation based on Ceará Education Board (SEDUC)

3. What is the impact of PEN10 on learning outcomes?

Program participation was determined by schools' rank[19] in the state-wide learning assessment, a design that allows for a causal estimate of program impact through a regression discontinuity design. This method compares schools that *barely* made it into the program with schools that *barely* didn't. Even if schools have some influence over their performance, principals are unable to *precisely* manipulate their rank to participate in the program. Therefore, very close to the participation cutoff, participation in the program is essentially as good as randomly assigned. This identification strategy provides better guidance of PEN10 impacts than simply comparing the outcomes of schools that participated in the program with schools that did not, which may give biased estimates of the program participation effects. This section reports the results from Goldemberg (2019), which used this empirical design, exploring the discontinuity at the participation cutoff, the 150th participant, both for top and bottom-performing schools.

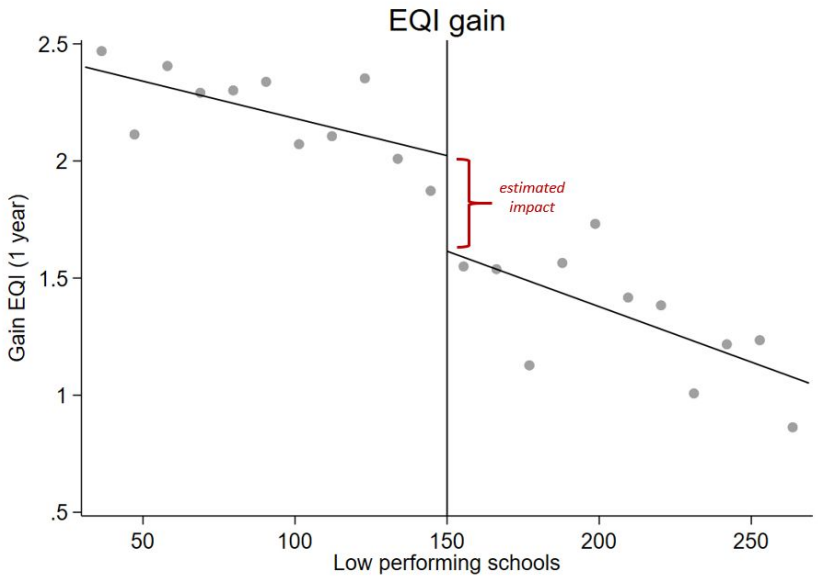
Administrative data on the peer-mentoring initiative was combined with publicly available education data to form a school-level panel of all public schools in the state. Between 2009 and 2015, the seven PEN10 editions for second graders results established a total of 1,048 partnerships among 1,429

[18] Infrastructure score refers to the average of five indicator variables: whether a school has a library, a sports court, a teachers' room, a science laboratory, and a computer laboratory.
[19] Rank in terms of the education quality index (EQI), with a set of eligibility and tie-breaker rules which evolved over the years. See Appendix 1 for a comprehensive description of the rules' evolution and their corresponding legislation.

unique schools based on their EQI, all over the state[20]. It is worth highlighting that 62 schools that had participated as low performers, after a successful turnaround, were selected to participate as top performers offering mentorship.

This evaluation shows that low-performing schools that received PEN10 mentoring improved learning outcomes without negatively affecting the top-performing schools serving as mentors. At the discontinuity, the effect of participating in the program for low-performing schools is an improvement of almost half a point in the EQI[21] after one year of mentoring (Figure 4). Program participation also increases by 15 percentage points the likelihood that previously struggling schools turn around successfully, surpassing the minimum bar of an EQI of 5. The schools that *barely* didn't make it into the program, which is the comparison group, also had an increase in their EQI as a result of the general improvements in Ceará. The selected empirical strategy nets out the change in the comparison group.

Figure 4 - PEN10 participation effects on low-performing schools



Notes: Change in EQI after one year, for low-performing schools that were eligible to participate, around the participation cutoff. Schools with rank≤150 were selected to join the program, while those with rank>150 were not. The jump at the rank 150 is the estimated causal impact of the program. Data presented in bins. For technical details, see Appendix 2.

The effects of program participation for top-performing schools, which receive a grant and provide mentorship, are not statistically significant. The program could potentially impact top-performing schools in different directions. On one hand, the immediate cash grant and the public recognition entailed by participation could boost staff extrinsic and intrinsic motivation and increase the availability of educational inputs. On the other hand, the extra work taken by the top principal, serving as a mentor, could harm student outcomes – despite the financial incentives that conditional the second disbursement to both own EQI maintenance and mentee school improvement. While there is no measure of the effects

[20] See Appendix 2 for more information on the data and empirical strategy used in the impact evaluation.
[21] The precise estimate is 0.484 (0.18 SD), significant at 5 percent, in our preferred specification. Regression tables including alternative specifications, varying bandwidth and other parameters are presented in the Appendix 2.

of the recognition and professional satisfaction gained by the top-performing participant schools, there is anecdotal evidence that they are motivated to engage as mentors[22]. The evaluation finds no significant EQI impact for the top-performing schools. The lack of statistically significant results indicates that the extra workload on principals in top schools did not affect their excellent performance. Since top-performers are already very close to the maximum EQI score, a ceiling effect may also contribute to this finding.

The huge gap between top- and bottom-performing schools may have facilitated the results. One important caveat is that each year the program selected only 150 pairs of schools from approximately 3,000 schools that satisfy the minimum enrollment criteria for participation. The gap between the top- and bottom-performing principals possibly plays a role in explaining the productive outcomes of the mentoring relationship and is likely much larger at the 150-cutoff. In that case, if the program was significantly expanded, results could be smaller, as paired schools could have less to learn from each other[23].

The average effect masks significant heterogeneity of performance gains across schools, likely driven by mentorship implementation and fit between mentor and mentee. Considering all the participants, thus abandoning causality claims, the average change for the low-performing schools is 2.4 points in EQI after a year, departing from a baseline of 2.8 (in a 0 to 10 scale). The results of the mentorship relationship for each established pair and their varying degree of success in improving student outcomes may be driven by various mechanisms. One factor could be mentorship implementation – whether the assigned principals did meet at least six times, and the quality of those mentoring sessions (i.e., whether they were live or online, their duration, a possible exchange between other members of the school leadership or teaching teams). Another important mechanism is possibly the fit between mentor and mentee (i.e., whether the mentor can understand the context of the mentee and provide efficient strategies, whether they have positive personal interactions). However, we do not observe data to further test these hypotheses on the frequency and quality of the meetings between mentor and mentee.

The similarity in the number of enrolled students and proximity between paired schools is associated with stronger outcomes for low-performing participants. There is a significant relationship between the proximity of matched schools and successful turnover. Belonging to the same district – which nests on average 9 neighboring municipalities – is associated with a 0.6 increase on the outcome of interest. Consistently, the average driving time between paired schools is 3 hours, with one additional hour being associated with a 0.1 lower gain in EQI. Being paired with a school located at a short distance seems more effective, possibly because those pairs succeed at implementing a more intense and better-quality mentorship relationship. Another important factor in determining partnership success is the similarity in school size. This heterogeneity has important policy implications for policy makers considering implementing PEN10.

---

[22] The award ceremony, when the participant schools sign the letter of intentions, is a highly celebrated event and an exceptional opportunity to engage all schools in the program activities.

[23] A regression discontinuity design estimates local average treatment effects (LATE). In this case, program participation effects were estimated at the 150 cutoff, which corresponds roughly to the bottom 5% of eligible schools. The further away from this cutoff, the less informative this estimate becomes of the potential impacts of the program.

While this evaluation provided convincing evidence of the positive impacts of the program, additional analysis could better illuminate its mechanisms. Areas for further exploration include the analysis of longer-term effects and impact on education equity. Moreover, a study of grants usage and mentorship meeting plans could provide a benchmark of best practices. Nevertheless, this remains an example of a successful peer-mentoring intervention for school principals in a developing context.

4. Implications for public policy

The peer-mentoring initiative for school principals is a crucial component of Ceará's reform to improve early-grade literacy and shows that a peer-mentoring initiative may be an affordable and relatively simple strategy to improve learning outcomes of low performing schools. PAIC has proven to be responsible, along with other reforms and programs, for a significant evolution in the state's education quality outcomes[24]. PEN10 is a core initiative in the PAIC program, taking over half of its budget. Therefore, assessing PEN10 effects on learning is indispensable to understand the mechanisms through which the reform has impacted Ceará's quality of education. In a policy perspective, a peer-mentoring initiative, based on knowledge-sharing, may be a less expensive strategy to provide management training for principals in schools with low performance than formal training or replacing those principals. This strategy fits the concept of self-improving school systems – policy contexts that emphasize competition and collaboration between schools as a means of achieving improvement[25]. The assumption is that without technical assistance from the top-performing schools' principals, the mentees would not have had access to the same knowledge to improve their managing skills in the short term.

The program helped to lift low-performance schools without harming the performance of the top schools. The evaluation findings suggest that school turnaround may be achieved systemically. Program participation improved the education quality index of the low-performing schools, through significant gains in student learning and equity outcomes in only one year. The program did not negatively affect student learning in the top-performing schools that provided mentoring. Program structure and mechanisms, such as the award ceremony, could have played an important role to motivate schools into the mentorship meetings.

Similar size and geographical proximity of matched schools are associated with larger mentorship's outcomes. Understanding the drivers and enabling factors for the program's success is key to expanding or replicating it in other contexts. The heterogeneity has important policy implications for other education districts considering replicating this program. We did not test an exhaustive combination of pairing types due to the lack of data, but, in this exercise, size and proximity stand out in terms of performance of paired schools. That is it: when the paired schools have a similar number of enrolled students and are

[24] Loureiro et al., 2020
[25] Ainscow, 2015

nearby to each other, the outcomes for low-performing participants are significantly higher. Therefore, peer-mentorship programs can take advantage of this result by having specific rules when pairing schools.

For instance, teacher qualification, measured by the share of teachers with a bachelor's degree, and school infrastructure are significantly worse in the top-performing schools, which points to the importance of other variables such as principals' ability. Though counterintuitive, the fact that low-performing schools are not worse endowed in terms of teachers nor physical infrastructure may be a condition for the program's positive outcomes. That is, the performance difference between the top and bottom schools may arise partly due to some unobserved principal's ability gap[26], such as managing skills, which may be transferable through mentoring. This study does not disentangle the impacts of the peer-mentorship from the cash grant received, a natural next step for future research.

The success of the peer-mentoring initiative is likely related to the other reforms and programs implemented in the state. At the same time the state of Ceará implemented PAIC and its peer-mentoring component PEN10, it also provided incentives for municipalities achieving better education results, through a results-based financing system. Other enabling conditions developed along with the reforms included giving autonomy to municipalities to design and implement education policies and establishing a reliable monitoring & evaluation system. This context is important to explain how Ceará put learning as the ultimate goal for the education reforms, constituting a favorable setting for PEN10 to succeed.

References

Ainscow, M. 2015. Towards self-improving school systems: Lessons from a city challenge. London: Routledge.

Baroody, Karen. 2011. "Turning around the Nation's Lowest-Performing Schools: Five Steps Districts Can Take to Improve Their Chances of Success." Center for American Progress. https://eric.ed.gov/?id=ED535862.

Calonico, S., M. D. Cattaneo, Mh Farrell, and R. Titiunik. 2017. "Rdrobust: Software for Regression-Discontinuity Designs." Stata Journal 17 (2): 372–404.

Costa, Leandro Oliveira, and Martin Carnoy. 2015. "The Effectiveness of an Early-Grade Literacy Intervention on the Cognitive Achievement of Brazilian Students." Educational Evaluation and Policy Analysis 37 (4): 567–90. https://doi.org/10.3102/0162373715571437.

Fouche, T. P., 2011. "Effects of the Texas Principal Excellence Program on Texas Principal Leadership Behavior and School Outcomes", Denton: University of North Texas.

[26] Considering that PAIC specified the technical criteria used by all municipalities in their principals' selection process, it is unlikely that the principal's ability gap is associated with differences in the selection processes in top- and low-performing schools.

Gates, S. M., Hamilton, L. S., Martorell, P., Burkhauser, S., Heaton, P., Pierson, A., Baird, M., Vuollo, M., Li, J., Lavery, D., Harvey, M., and Gu, K.. 2014. “Preparing Principals to Raise Student Achievement: Implementation and Effects of the New Leaders Program in Ten Districts”, Santa Monica, Calif.: RAND Corporation, RR-507-NL, 2014. http://www.rand.org/pubs/research_reports/RR507.html

Goldemberg, Diana 2019. “Essays on Economics of Education.” PhD diss., EPGE/FGV, Rio de Janeiro. http://hdl.handle.net/10438/27359

Herman, Rebecca. 2012. “Scaling School Turnaround.” Journal of Education for Students Placed at Risk (JESPAR) 17 (1–2): 25–33. https://doi.org/10.1080/10824669.2012.637166.

Herman, Rebecca, Priscilla Dawson, Thomas Dee, Jay Greene, Rebecca Maynard, Sam Redding, and Marlene Darwin. 2008. “Turning around Chronically Low-Performing Schools: A Practice Guide | Center for Education Policy Analysis.” Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=7.

Herman, R., Gates, S.M., Arifkhanova, A., Barrett, M., Bega, A., Chavez-Herrerias, E., Han, E., Harris, M., Migacheva, K., Ross, R., Leschitz, J., and Wrabel, S. 2017. “School Leadership Interventions Under the Every Student Succeeds Act: Evidence Review: Updated and Expanded”. Santa Monica, CA: RAND Corporation.

IBGE (2017). Pesquisa nacional por amostragem de domicílios. Retrieved October 20, 2018, from http://www.ibge.gov.br/

INEP (2018). Censo Escolar 2007 - 2017 and Prova Brasil 2017. Retrieved October 20, 2018, from http://inep.gov.br/web/guest/microdados

Jacob, R., R. Goddard, M. Kim, R. Miller, and Y. Goddard, 2015. “Exploring the Causal Impact of the McREL Balanced Leadership Program on Leadership, Principal Efficacy, Instructional Climate, Educator Turnover, and Student Achievement,” Educational Evaluation and Policy Analysis, Vol. 37, No. 3.

Loureiro, Andre, Louisee Cruz, Ildo Lautharte, and David Evans. 2020. "The State of Ceará in Brazil is a Role Model for Reducing Learning Poverty".

Meyers, Coby, and Jonathan Sadler. 2018. “District Leaders Engaged in School Turnaround: Identified Challenges and Espoused Responses.” NASSP Bulletin 102 (2): 89–110. https://doi.org/10.1177/0192636518786008.

Nunnery, J. A., S. M. Ross, S. Chappell, S. Pribesh, and E. Hoag-Carhart, 2011. “The Impact of the NISL Executive Development Program on School Performance in Massachusetts: Cohort 2 Result”, Norfolk, Va.: Center for Educational Partnerships at Old Dominion University.

Riley, Derek L., and Julie Meredith. 2017. “State Efforts to Strengthen School Leadership: Insights from CCSSO Action Groups.” Policy Studies Associates. https://www.wallacefoundation.org/knowledge-center/pages/state-efforts-to-strengthen-school-leadership.aspx.

Service, Brenda, Gulay Erin Dalgic, and Kate Thornton. 2018. “Benefits of a Shadowing/Mentoring Intervention for New Zealand School Principals.” Professional Development in Education 44 (4): 507–20. https://doi.org/10.1080/19415257.2017.1378705.

Spiro, Jody, Mary C. Mattis, and Lee D. Mitgang. 2007. “Getting Principal Mentoring Right: Lessons from the Field.” Perspective. The Wallace Foundation. https://www.wallacefoundation.org/knowledge-center/pages/getting-principal-mentoring-right.aspx.

Sternberg, R. J., Birney, D., Kirlik, A., Stemler, S., Jarvin, L., and Grigorenko, E. L. 2006. “From Molehill to Mountain: The Process of Scaling Up Educational Interventions (Firsthand Experience Upscaling the Theory of Successful Intelligence).” In Translating Theory and Research into Educational Practice: Developments in Content Domains, Large-Scale Reform, and Intellectual Capacity. The Educational Psychology Series. Mahwah, N.J.: Lawrence Erlbaum Associates.

Trujillo, Tina, and Michelle Renée. 2015. “Irrational Exuberance for Market-Based Reform: How Federal Turnaround Policies Thwart Democratic Schooling.” Teachers College Record 117 (6).

Appendix 1 - List of the state legislation on PEN10

This is an exhaustive list of the state legislation creating and regimenting PEN10:

- **Law nº 14.371 / June 19, 2009**: creates PEN10. Defines the number of participants as 150 schools in each category (*premiada & apoiada*). Eligibility conditioned on at least 20 enrolled second graders and participation rate above 50% on SPAECE. Additionally, top performers must score above 8.5 in the education quality index (*Índice de Desempenho Escolar na Alfabetização*, IDE-Alfa). Grants for top performers are set to R$2,500 per enrolled grade 2 student, paid 75% in the first disbursement. Low performers receive R$1,250 per enrolled grade 2 student, paid 50% in the first disbursement
- **Decree nº 29.896 / September 16, 2009**: defines rules for the second, conditional disbursement, and provide guidelines for the grant use
- **Law nº 14.580 / December 21, 2009**: specifies that the minimum enrollment of 20 students condition has SPAECE application date as the reference point
- **Law nº 14.949 / June 27, 2011**: expands the program to the grade 5 assessment
- **Law nº 15.052 / December 6, 2011**: increases minimum participation rate of students in SPAECE to 90%, establishes tie-breaking rules and a new eligibility condition (only schools in municipalities with above 70% of students with satisfactory achievement are eligible as top performers), limits repeated participation
- **Decree nº 30.797 / December 29, 2011**: reflects a change in the calculation method of the education quality index (IDE-Alfa) and regiments the grade 5 program (based on IDE-5)
- **Law nº 15.246 / December 6, 2012**: modifies tie-breaking rules and imposes additional eligibility condition (only schools in districts above 70% of students with satisfactory achievement are eligible as top performers)
- **Law nº 15.923 / December 15, 2015**: expands the program to the grade 9 assessment and modifies the grade 2 version to discontinue the peer-mentoring, preserving only the prize for top performers. The selection of low performers to be matched with top performers remains for grade 5 and grade 9 levels. Grants values for top performers changed to R$2,000 per enrolled student and R$1,000 for low performers
- **Decree nº 32.079 / November 9, 2016**: regiments the grade 9 program (based on IDE-9) and revises guidelines for grant disbursement and use

Appendix 2 - Data and Empirical Strategy for the Impact Evaluation

The state education board provided administrative data reporting which partnerships were formed each year and how much was disbursed to each participating school. Between 2009 and 2015[27], the seven PEN10 editions for second graders results established a total of 1,048 partnerships[28] among 1,429 unique schools, all over the state. It is worth highlighting that 62 schools that had participated as low performers, after a successful turnaround, were selected to participate as top performers offering mentorship.

The school ranking is based on their EQI, Education Quality Index, calculated as follows for each school at every round of the annual state-wide assessment of second graders (SPAECE Alfa):

$$EQI = LP * PR * EAF$$

where LP is the average reading proficiency measured in a scale from 0 to 10 points, PR is the participation rate, the share of enrolled second graders assessed, varying from 0 to 100%, and EAF is an equity adjustment factor, also varying from 0 to 100%. The EAF is calculated as a weighted average of share of students in each reading proficiency level: $EQI = 0 * SL_1 + 0.25 * SL_2 + 0.50 * SL_3 + 0.75 * SL_4 + 1 * SL_5$, where SL_1 is the share of students in the lowest level of proficiency and SL_5 is the share of students in the highest level of proficiency.

Both the EQI and its three components (average reading proficiency, participation rates, and equity adjustment factor) are publicly disclosed. We use the SPAECE Alfa results from 2008 to 2017 to construct a panel at the school level, in which we can observe up to three years of lagged performance from program participation. Relevant school characteristics, obtained from Brazilian School Census by INEP (National Institute for Educational Studies and Research), were also included in the analysis.

Program eligibility is conditional on having at least 20 students enrolled in second grade. Thus, only 44% of Ceará's public elementary schools were ever eligible to participate in the program, but those schools are significantly larger, representing 70% of the enrollment in the state, and more urban. Table 1 presents descriptive statistics for all public elementary schools in Ceará (column 1) and two subsamples of interest: low-performing schools that are just below and just above the participation cutoff of rank 150 (columns 3 and 4). By design, the baseline EQI of those barely included to participate is significantly lower, though it is interesting to note that those two groups do not differ in participation rates in the state-wide assessment. Column 5 reports the p-value for equality of means between the groups. Column 6 reports the p-value for a discontinuity in the baseline characteristics at the participation cutoff for low-performing schools.

[27] Because Ceará had been so successful in improving early-grade literacy, in 2015 it changed focus to promote educational improvement in later grades. *Prêmio Escola Nota Dez* selected up to 150 pairs of schools based on 2nd-grade assessment from 2008 to 2014, plus up to another 150 pairs based on 5th-grade assessment since 2011 and up to 150 pairs based on 9th-grade assessment since 2015. Despite discontinuing the mentorship component based on 2nd-grade results and the financial incentives for low-performing schools, the program continues to award the 150 top performers in this grade-level.

[28] In the first year, only 148 partnerships were established. The remaining six years had 150 partnerships each.

Table 1 - Descriptive Statistics (pooled 2009-2015)

	(1) All public elementary schools	(2) All elegible schools in Ceará	(3) Top performing participants	(4) Low performing participants	(5) Low performing "barely" included (101th-150th)	(6) Low performing "barely" excluded (151st-200th)	(7) P-value equal means col. 5 & 6	(8) P-value RDD col. 5 & 6
Sample size								
Number of Unique Schools	6,676	2,934	704	787	317	313	-	-
Number of Observations (school-year)	33,476	12,223	1,048	1,048	350	350	-	-
School Characteristics								
Share urban	0.33	0.68	0.64	0.71	0.72	0.74	0.58	0.32
Infrastructure score	1.62	2.68	2.51	2.88	2.97	2.99	0.89	0.67
Students								
Number of 2nd graders per school	24.7	47.6	38.8	52.1	54.8	56.1	0.68	1.00
Share of students using school transpc	0.27	0.24	0.28	0.18	0.18	0.19	0.67	0.97
Share of students behind ideal grade	0.16	0.17	0.09	0.23	0.23	0.22	0.53	0.05
Teachers								
Number of teachers per school	7.9	12.7	11.6	13.3	13.1	14.5	0.03	0.59
Teacher average age	36.6	37.7	36.1	39.1	39.0	39.0	0.98	0.95
Share of female teachers	0.85	0.86	0.86	0.84	0.84	0.84	0.67	0.92
Share of teachers with B.A.	0.60	0.73	0.67	0.76	0.77	0.75	0.27	0.98
Baseline Education Quality Index								
Participation rate on assessment	0.95	0.96	0.99	0.91	0.92	0.93	0.95	0.91
Literacy proficiency	7.52	7.83	9.94	5.63	5.95	6.10	0.04	0.94
Equity adjust factor	0.73	0.77	0.98	0.49	0.55	0.57	0.05	0.92
Education Quality Index	5.72	6.14	9.67	2.77	3.21	3.43	0.04	0.94

This table shows descriptive statistics for all public schools in Ceará which offer second grade (1), schools large enough to participate in the program (2), the top performing participants (3), the low performing participants, the subsample of elegible low performing school just below the participation cutoff (5) and the subsample of elegible low performing schools just above the participation cutoff (6). Column 7 tests for equality of means and Column 8 for a discontinuity in baseline characteristics at the participation cutoff for low performers (using Equation 1), and both report the corresponding p-values. Note that the running variable is the rank of EQI, thus, one should expect that EQI and its components would have significantly different means between (5) and (6), as shown in Column 7, but no discontinuity, as captured by Column 8.

Source: Goldemberg (2019)

Main specification

The main specification is a linear regression for eligible schools ranked *around* 150th, where *around* is defined according to the optimal bandwidth selection of Calonico et al. (2017). The following equation estimates the effects of participating in the program at both ends – separately for the top and low performer schools[29]:

$$Y_{sc(y+1)} = \alpha_c + \beta_c 1\ \{r_{scy} \leq 150\} + \sigma_c r_{scy} + \gamma_c\ 1\ \{r_{scy} \leq 150\}\ r_{scy} + X_{sy}'\delta + \varepsilon_{scy} \qquad (1)$$

The running variable is the rank of a given school amongst all eligible schools in the low/top category of participation in a year (r_{scy}). The treatment variable is $1\ \{r_{scy} \leq 150\}$, an indicator variable equal to one if the school was ranked below 150 amongst eligible low/top performers and, hence, selected to participate in the mentorship program. Our main parameter of interest is β_c, the effect of participation in the program for schools in the cutoff, estimated for the top and low categories of participation. The outcome variable $Y_{sc(y+1)}$ reflects the temporal evolution of the school's EQI, either as the observed change relative to the baseline ($EQI_{sc(y+1)} - EQI_{scy}$) or as an indicator variable reflecting the achievement of the

[29] For Equation (1) to estimate the causal effects of program participation, the key identification assumption is that any potential factors that influence the outcomes are continuous around the cutoff rscy=150, and, thus, any discontinuity in outcomes at the cutoff is the result of the program. That is, for any potential factor X that influences the outcome Y, such as Y=f(r,X), f(X) needs to be continuous, with only r being discontinuous at the threshold. By design, our running variable has a uniform distribution, precluding the need for any formal test for manipulation of the running variable. Further evidence that lends support to our identification assumption is that we do not find evidence of discontinuity in covariates at the program cutoff, as reported in Table 1, column 6 for the low-performing schools.

target performance of at 5 for the low-performing schools $1\left\{EQI_{sc(y+1)} \geq 5\right\}$ and the maintenance of good results for the top-performing schools $1\{EQI_{sc(y+1)} \geq EQI_{scy}\}$. We include as controls X_{sy} the baseline EQI and a year dummy to control for a general time trend between editions of the program. Standard errors are clustered at the school level. The estimates consider the pooled data from the seven years of the program, 2009-2015, and robustness checks are performed. Table 2 presents the results for low-performance schools.

Table 2 - Regression Discontinuity Design (low performers, 2009-2015)

	Low Performers						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Delta EQI							
Participation effect	0.484	0.433	0.595	0.265	0.466	0.452	0.445
	(0.218)**	(0.197)**	(0.229)***	(0.311)	(0.251)*	(0.196)**	(0.181)**
Using bandwidth	83	90	81	41	62	104	124
(B) Goal reached *(EQI>=5)*							
Participation effect	0.154	0.152	0.153	0.094	0.123	0.161	0.147
	(0.061)**	(0.057)***	(0.086)*	(0.091)	(0.073)*	(0.054)***	(0.049)***
Using bandwidth	69	70	67	34	52	86	103
Specification							
Bandwidth method	Opt	Opt	Opt	.5*Opt	.75*Opt	1.25*Opt	1.5*Opt
Year dummies	No	Yes	No	No	No	No	No
Analytical weights	No	No	Yes	No	No	No	No
Observations	12,205	12,205	12,205	12,205	12,205	12,205	12,205

Notes: Coefficients on program participation for low performing schools from regressing each of the outcomes on the running variable of the RDD (rank), program participation (rank<150) and the interaction of these two variables for the set of schools with |rank-150|<Using bandwidth. Outcomes are the change in EQI after one year (Panel A), or an indicator variable for whether the school reached the EQI goal of 5 (Panel B). Heteroskedasticity robust standard errors clustered at the school-level are in parentheses (p<.10 ** p<.05*** p<.01).*

Source: Goldemberg (2019)

Heterogeneity analysis

To investigate the drivers of turnover success, we regress the observed improvement of the low performer participants on some characteristics compared to their top performer counterparts. The coefficients estimated in this manner do not allow for any causal interpretation and should be interpreted as correlations. Nevertheless, this exercise is useful to provide indicators of what may constitute a successful mentorship between two school leaderships. We use the following equation:

$$Y_{s_t s_l y} = \alpha_y + X_{s_t s_l}'\beta + \varepsilon_{s_t s_l y} \qquad (2)$$

where $Y_{s_t s_l y}$ is the outcome obtained by the mentee low-performing school s_l, from partnering with the mentor top-performing school s_t, after participating in the program in year y. We focus on the evolution of the EQI of a school relative to the baseline ($Y_{s_t s_l y}$). We model this outcome as a linear regression with constants by year of partnership α_y and an explanatory variable $X_{s_t s_l}$ capturing some aspect of interest in the partnership between the two schools. Our parameter of interest β captures the relationship between the

explanatory variable and the obtained outcome. We use as $X_{s_t s_l}$ two indicator variables (i) whether both schools are located in rural or in urban zone; (ii) whether both schools belong to the same education district; and three continuous variables: (iii) the dissimilarity of school sizes, which takes the value 0 if both schools have the same number of enrolled second graders, according to $abs(n_{s_t} - n_{s_l})/max(n_{s_t}, n_{s_l})$, (iv) the driving time in hours between the schools, and (v) the absolute difference in school socioeconomic status[30].

Table 3 reports the estimation results from equation (2), for each *X* variable individually first, then jointly estimated. Though not causal, they suggest a strong relationship between the similarity of matched schools and successful turnover. Belonging to the same district – which nests on average 9 neighboring municipalities – is associated with a 0.63 increase on the outcome of interest. Consistently, we observe a negative relationship between driving time and EQI delta: a one-hour increase in the travel time between schools is associated with a 0.07 lower delta. Those results suggest that being paired with a school located at a short distance is more effective – perhaps because those pairs succeed at implementing a more intense and better-quality mentorship relationship. Other important factors in determining partnership success are the similarity in school size and in socioeconomic status. The indicator variable of both schools being urban or rural is not significant.

Table 3 - Performance improvement regressions on mentorship characteristics

	Descriptive	**Coefficient on Low Performer's EQI delta**					
	Mean	(1)	(2)	(3)	(4)	(5)	(6)
Relational dummy variables							
X1: Same type (urban or rural)	0.56	-0.13					-0.07
	(0.50)	(0.12)					(0.13)
X2: Same education district	0.26		0.63				0.35
	(0.44)		(0.14)***				(0.18)**
Relational continuous variables							
X3: Dissimilarity of school size	0.40			-1.41			-1.19
	(0.25)			(0.22)***			(0.27)***
X4: Driving time (hours)	3.05				-0.07		0.01
	(1.94)				(0.03)**		(0.04)
X5: Absolute difference in SES (pp)	3.71					-0.07	-0.04
	(2.65)					(0.02)**	(0.03)
N (partnerships)		1048	1048	1048	1048	816	816

Note: left panel reports the mean value and standard deviation of the variable in the dataset. Right panel reporsts coefficients of regressions on EQI delta by low performers that participated in the program on the selected characteristic of the matched school pairs. All regressions include year fixed-effects. Heteroskedasticity robust standard errors clustered at the school-level are in parentheses (p<.10 ** p<.05*** p<.01).*

Source: Goldemberg (2019)

[30] We use the socioeconomic status (SES) index calculated by INEP for the Brazilian schools. INEP uses several socioeconomic questionnaires of nation-wide assessments to calculate the SES index. The items used are related to family income, the ownership of goods, and the parents' educational level. Due to missing data or a small number of students in part of the schools, INEP does not calculate the index for all schools.