Heliyon 9 (2023) e15130

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

An analysis of student numbers per class and educational performance in the Brazilian context

André Calixto Gonçalves *, Leandro Campi Prearo, Nonato Assis de Miranda, Paulo Roberto Mendes

Municipal University of São Caetano do Sul, 3400 - Goiás Avenue, Stricto Sensu - Barcelona Campus Center, São Paulo, Brazil

ARTICLE INFO

Keywords:
Efficiency
Educational economics
Resource allocation
Economic impact
Propensity score matching
Quantitative methods

ABSTRACT

The relationship between class size and school performance has always been ambiguous and the current literature has found no direct connection between them, especially in the Brazilian context. Therefore, this study aimed to verify whether the number of students per class influences school performance. We used Microdata from the Prova Brazil of 2017. Using the propensity score matching statistical model, with the nearest neighbor matching estimator, we grouped the classes into clusters by similarity. The metric used to group the clusters was the Euclidean distance. We attempted to verify adherence to the normal distribution of data using the Kolmogorov Smirnov test and tested the null hypothesis of the medians using the Wilcoxon test. All the statistical analysis were performed using SPSS Statistic version 20. The results showed that the number of students per class has little influence on performance and, when the influence exists, larger classes perform better.

1. Introduction

Investigations of the determinants of student performance have guided much of the research in the field of economics of education in different countries. Although the literature is not recent, there is no consensus on what are the main inputs of the educational production function, let alone on their real effects [1–10, 11, 12].

Class size can be an extremely appealing policy instrument, as it is often regarded as a relatively straightforward measure for improving school performance. This is because it is enough to have the resources to increase the number of rooms and hire more teachers or increase the size of the room and provide the necessary seats [13].

There are several studies on the increase or decrease in class size. The Star Project, carried out in Tennessee, demonstrated that smaller classes obtained better results in terms of learning and academic performance. However, class size reduction carried out in California did not yield any significant result [1,2]. All these studies contribute to literature the actual influence of class size on performance.

Among other aspects, for example, there is no consensus on class size and its influence on student performance. In general, common sense among students, parents, teachers and administrators argues that smaller classes generate better results in terms of performance. This understanding is based on the assumption that classes with fewer students allow the teacher to be better able to identify the

* Corresponding author.
E-mail addresses: andre.calixto.goncalves@gmail.com (A.C. Gonçalves), leandro.prearo@gmail.com (L.C. Prearo), mirandanonato@uol.com.br (N.A. Miranda), paromendes@gmail.com (P.R. Mendes).

https://doi.org/10.1016/j.heliyon.2023.e15130
Received 11 May 2022; Received in revised form 17 February 2023; Accepted 27 March 2023
Available online 8 April 2023

individual needs of students. On the other hand, in larger classes, students tend to be more dispersed, which can generate disciplinary problems, consuming a good part of the class, thus generating learning losses for students.

However, these conceptions do not find support in the empirical literature. Some studies have found a positive effect of the policy of reducing classes on student performance, others contrast these results showing a small or even non-existent impact. Like this study, in turn, other studies highlight that more students in the classroom can improve school performance.

In view of the above, this investigation was guided by the following question: what is the relationship between the number of students per class and student performance?

The aim of this study is to determine whether a similar relationship between class size and performance can exists in Brazilian schools. To that end, 16,607 Brazilian schools were evaluated, with 17,533 classes in the 5th grade (Elementary School) and 13,608 classes in the 9th grade (Middle School). The subjects evaluated were Portuguese language, with focus on reading, and mathematics focused on problem-solving.

2. The Brazilian context

The basic education system in Brazil is based on the 11 principles defined by the Law of Guidelines and Bases of Education [14]:

I. Equal conditions to access and stay at school
II. Freedom to learn, teach, research, and disseminate culture, art, and knowledge
III. Pluralism of pedagogical ideas and conceptions
IV. Respect for freedom and appreciation for tolerance
V. Coexistence of public and private educational institutions
VI. Free public education in official establishments
VII. Valuing the professional in school education
VIII. Democratic management of public education in the form of the law and legislation of education systems
IX. Quality standard assurance
X. Valuing extra-school experience
XI. Link between school education, work, and social practices.

In line with these principles, Brazilian school education is composed of two levels: the first, basic education, comprises of early childhood and elementary education, and the second comprises high school (Government of Brazil, 1996).

In early childhood education, the first phase involves daycare, where the child develops up to the age of three years and 11 months. The second phase involves preschool for two years aimed at the integral development of children up to five years of age [15]. After early childhood education, children undergo compulsory and free elementary education at public schools for 9 years divided into two phases—the first 5 years include children aged 6 to 10 and the remaining four years include children aged 11 to 14 (Government of Brazil, 2016). Finally, children attend high school with a minimum duration of 3 years. Aimed at children aged 15 to 17, it represents the last stage necessary for the completion of basic education in Brazil (Government of Brazil, 2016). The Brazilian government encourages educational development of Brazilian teachers working in basic education. As of 2013, the legislation had determined that all teachers working in all stages of the structure had to have at least a higher level in a degree course or full graduation from universities and higher education institutes (Government of [16].

According to data from Ref. [17] School Census, the number of students enrolled in basic education is approximately 47.3 million, distributed across 179,533 schools, of which 81.4% and 18.6% study in public and private education respectively. Out of 47.3 million students, 26.7 million are in elementary school, 7.6 million are in high school, and 13 million are in kindergarten (Government of Brazil, 2020). The number of teachers working in basic education in Brazil is 2.2 million, of which 63% work in elementary education (Government of Brazil, 2020).

According to the OECD, Brazil has developed since the 2000s in terms of enrollments and quality of education, with less favored students or those students who have low socioeconomic status reaping improved benefits [18]. The OECD also reveals that Brazil dedicates a relatively large portion of its Gross Domestic Product - GDP to education. The investments are largely decentralized, and this can contribute to reducing inequalities in education [18]. The budget execution report by the Brazilian Ministry of Education reveals that the expenditure per student in 2021 on basic education was US $747 on average—a value that is far lesser than that of countries like the USA, Finland, and China for example (Government of Brazil, 2018).

In terms of class size, the indicators of the Brazilian Ministry of Education for the year 2020 show that the country has an average of 18.3 students per class in the initial grades and 23.2 students per class in the final grades in the private school system. In the public school system, the averages are higher, with the initial grades having an average of 25.3 students per class and the final grades with 28.8 students per class (Government of Brazil, 2020).

To measure the quality of the Brazilian educational system, the government has large-scale assessments, the most important of which is the Prova Brazil (Government of Brazil, 2020). It is a large-scale diagnostic assessment in which tests are applied in the 5th and 9th grades of elementary school. The test evaluates the Portuguese language with a focus on reading, and Mathematics with a focus on problem-solving. In addition, the students complete a socioeconomic questionnaire (Government of Brazil, 2020).

One of the positive aspects of this evaluation is that teachers and principal of the evaluated schools also participate in the evaluation and answer questionnaires that collect data about their professional profile, working conditions, and school infrastructure (Government of Brazil, 2020).

In Brazil, the school principal is responsible for the division of classes; as the country does not have a general legislation on how this division should be done; Prova Brazil, through the principal's questionnaire, exposes the main methodologies used (Government of [19]: i) Homogeneity regarding age (students of the same age); ii) Homogeneity in terms of school performance (students with similar performance); iii) Age heterogeneity (pupils of different ages); iv) Heterogeneity in terms of school performance (students with different incomes); and v) Without discretion, or randomly (Government of Brazil, 2017).

Therefore, a survey that only uses data from schools that form classes "without criteria or randomly" can prevent a possible pre-selection previously applied by the school principal from damaging the data. Thus, the use of this criterion in this study made it possible to avoid the contamination of results that a pre-selection previously made by the principal could cause.

3. Literature review

When diagnosing the literature on the control of students by classroom to measure academic performance, some surveys, such as the STAR Project in Tennessee, USA stand out; it was conducted between 1985 and 1989 to evaluate the performance of students during their school development [1]. According to data from the project's press report published in 1999, the study carried out in the STAR project monitored the performance of a group of students placed from the third grade of preschool (kindergarten 3) in small classes and another group of students from regular classes [1].

The comparison of these 2 groups indicated that students in small classes performed better than those in large ones. Both improved learning and academic development were observed in the students from small classes and teachers reported improvements in the teaching dynamics [1]. [20] also found that the STAR project is the most influential and reliable study on reducing the class size and, therefore, several studies use it as a basis. Other prominent surveys that can be found in the class size literature are Wisconsin's SAGE Program and the California Class Size Reduction Program [2,21]. As presented in the 2015 Final Report, the Wisconsin SAGE Program in the US was based on an analysis of the performance of students from kindergarten through 3rd grade. It was defined by the study that participating classes should have up to 18 students and be compared with regular non-participating classes. Thus, the comparisons were made to measure the performance of these groups in Mathematics and Reading [21]. The SAGE Program found positive and significant effects in reading when verified in kindergarten, but this was not repeated when mathematics was verified. Further, a lasting impact review revealed that students participating in the project were more likely to stay in school [21].

Although both studies (STAR and SAGE) are important to the literature, but the California Class Size Reduction (CSR) Program overlaps the other studies in terms of size and investment [1,21]. According to the Executive Summary of the year 2002, the California CSR (1996–2002) invested 1 billion US dollars in the first year alone and represented a historic milestone in world education [2].

The California CSR encouraged school districts to downsize their class sizes to 20 students or less, and as a reward awarded $650 per student in that condition. Although the project was implemented quickly, the relationship between size and performance was inconclusive; that is, although the performance on the Stanford Achievement Test (SAT) increased, it could not be linked to the decrease in class size [2].

Furthermore, as [13] pointed out, it can be challenging to find a significant outcome in a project with an annual expenditure of 1 billion dollars because it presents results with political bias rather than those that purely indicate the influence of class size on performance.

The studies conducted in the Netherlands by Ref. [22]; Hong Kong by Ref. [23]; and France by Ref. [24] are also significant. Schools in the Netherlands had classes in which the class size varied significantly [22]; randomly selected 416 schools, to verify the relationship between performance and class size for the 2nd, 4th, 6th, and 8th grades. According to the author, the findings were unexpected: students in larger classes performed better than those in smaller classes [22].

In Hong Kong, Galton and Pell (2012) demonstrated in their study that there was a decrease in students per class in 37 primary schools. The average number of students in these schools before the study were 38 students per class, but this had reduced to a minimum of 20 to a maximum of 25 students per class. The performance in the subjects of Chinese, English, and Mathematics, observed over 3 years in Primary 1 to Primary 3, also did not show any significant differences. They only found differences in the individual attention span between students and teachers [23].

In contrast, the study in France by Ref. [24] had a total of 100 classes reduced to an average of 10.5 students per class. Another 100 classes with an average of 21.3 students were maintained to form a basis for comparison. To assess changes in performance, students' performances were monitored for 2 years. Authors observed that class reduction positively affected performance only in the first year; in the second year, students from smaller classes returned to larger classes and this effect disappeared [24].

4. Database

The database used for this research is from the School Performance Assessment or Prova Brazil for 2017 which measures the performance of students from the final grades of elementary school and middle school in the Portuguese language focused on reading and mathematics focused on problem-solving (Government of Brazil, 2017).

The test is held every 2 years, between October and November. In 2017, Prova Brazil had approximately 70,000 schools and approximately 5.4 million students (Government of Brazil, 2017). To measure students' performance in the Portuguese language, Prova Brazil assigns a maximum of 350 and 400 points for students in the initial and final grades respectively; in mathematics, it assigns a maximum of 400 points and 450 points for students in the initial and final grades respectively (Government of Brazil, 2017).

In addition, Prova Brazil uses 4 questionnaires focusing on the characteristics of the school, principal, teacher, and student. The school's questionnaire has 74 questions focused on infrastructure, safety, and conditions of use. The principal's questionnaire has 111

questions focusing on their education, professional experience, functional characteristics, characteristics of the school team, policies, actions, programs, lunches, difficulties in school management, financial resources, textbooks, violence at school, and religious education (Government of Brazil, 2017).

The teacher's questionnaire has 125 questions focusing on their education, professional experience, functional characteristics, professional development, reading and cultural habits, and use of audiovisual and educational resources. In addition, there are questions about the integration of the school team, learning problems, school violence, didactics, use of time, pedagogical practices, and teacher expectations (Government of Brazil, 2017).

There are 2 students' questionnaires—one for each of the final grade students of elementary school and one for each of the final grade students of middle school, with 51 and 57 questions, respectively. Both questionnaires focus on student characteristics and habits, their socioeconomic status and that of their family, and family encouragement (Government of Brazil, 2017).

All data of the questionnaires can be cross-checked using unique codes, such as the unique code of the school, teacher, student, class, municipality, state, etc. Thus, by separating the questions that meet the criteria pre-defined by the survey, it is possible to carry out a complete study by cross-examining all the data.

5. Confounding variables

To avoid contamination by variables that directly influence performance, it is necessary to control for these variables. According to Ref. [25]; confounding variables can influence and confuse the results. They explain that these potentially confounding variables can influence what and how participants learn, and thus directly influence performance. Therefore, to control the variables that influence the performance of Brazilian students, some variables were chosen in the questionnaires based on the literature on performance determinants. [26–29, 30]. All variables have equal value in this study. These chosen variables were used after the questionnaires were answered (Table 1).

6. Method

In view of the objectives of this investigation, a quantitative research method was chosen. This method, usually presents empirical data in the form of numbers or in the form of measurements [31]. To accomplish this, the following analytical tools were used: Propensity Score Matching (PSM), Clustering, Euclidean Distance and Nearest Neighbor Matching (NNM).

We chose PSM because it is a tool for causal inference in non-randomized characteristic studies that allows the observation of large sets of intervening variables, as was our case. This tool allows controlling for different intervening variables, so that they do not influence the result to be found. Thus, when using PSM, a large set of intervening variables can be analyzed at the same time [32]. This statistical model can also adjust the effect of a treatment on intervening variables that could contaminate the result. In addition, the model is a commonly used alternative for regression fitting.

The data were then grouped using clustering analysis. According to [33], clustering is a very useful and basically ubiquitous tool in

Table 1
Confounding variables in Brazilian classes.

Students Variables	Teacher Variables	Principal Variables	School Variables
Sex	Sex	Sex	Existence of computers for students and teachers
Race	Age	Age	Internet access for students and teachers
Age	Undergraduate and graduate level	Undergraduate and graduate level	Existence of a library
Home computer	Number of jobs in addition to teaching	Number of jobs in addition to principal	Existence of sports court
Lives with the mother	Length of time as a teacher	Length of time as a principal	Existence of laboratories
Mother's education	Length of time as a teacher at this school	Participated in the last 2 years of professional development activity	
Mother can read or write	Number of schools in which it operates	There is a school council in this school	
Lives with the father	Participated in the last 2 years of courses in education	Class formation criteria	
Father's education	Participated in the last 2 years of specialization courses		
Father can read or write			
Parent attendance at meetings			
Family incentive to study			
Family encouragement to do homework			
Family encouragement to read			
Family incentive to go to school			
Parental attention to school situation			
Frequency of use of the library			
Economic class			

the analysis of statistical data. These authors state that clustering constitutes a grouping of data under the definition of similarity, where similar items are in the same group and different items are in different groups. By grouping the classes into clusters by degree of similarity, it is possible to compare them with each other. Thus, classes with approximately the same incidence of intervening variables were compared, preventing the contamination of intervening variables from harming the result.

[34] states that clustering is the division of data into groups of similar objects, in which some details are disregarded in favor of simplifying the data. The author explains that clustering data is a data modeling technique that provides concise, grouped summaries.

Mishra and Guha [33] understand that the most important aspect of cluster analysis is the definition of what constitutes a cluster, so as to verify why one is better than the other. The authors warn that the quality of the cluster is dependent on an in-depth analysis of all the points assigned to it. When defining the limit of the constitution of clusters, it is possible to define which classes make up each one of them, according to their propensity scores.

When grouping the classes into clusters, we noticed that some of them did not fit into any cluster, that is, they were not approximately similar to any other within the analyzed universe. These are the so-called multivariate outliers and since they cannot be used for comparison purposes, they had to be excluded. Additionally, because the number of classes in the sample was too large, it was decided that these classes should be grouped using the k-means clustering technique of [35,36]. But before applying the technique to group the classes by degree of similarity, it was also necessary to determine the limit of this similarity or distance. So we chose to use the Euclidean distance technique $\sqrt{((x1 - x2)^2 + (y1 - y2)^2)}$ by Ref. [37].

The Euclidean distance is a standard metric in geometric problems [38] that is widely used in clustering. It is the normal distance between two points and can be easily measured with a ruler in the two or three dimensions.

Once the classes had been grouped by degree of similarity and all the intervening variables were also controlled. It's was possible to ensure, each class would have a specific amount of incidence of these intervening variables and, for this reason, it was possible to group them into clusters by degree of similarity, defined by the use of the Euclidean distance and selected using the NNM technique of [39].

Based on Thoemmes (2012), after estimating the PSM, we started the correspondence between the scores obtained. This correspondence can be accomplished in many ways. The most commonly used is the NNM technique. The NNM technique was used in this study to select approximately similar participants within the measured database. To ensure good matches, it was also necessary to define the maximum difference allowed by the Euclidean distance.

The NNM estimation technique made it possible to assess whether students belonging to a class size of 20 students had the same observable characteristics, required to hypothetically determine the influence of school results and class size, as students belonging to a class size of 30.

In terms of procedures, we carried out the following steps:

First, all the data were filtered according to the criteria of class formation per the question presented in the principal's questionnaire to avoid a pre-treatment selection bias. Therefore, 16,607 schools with 17,533 and 13,608 classes in the initial and final grades respectively, were identified by only selecting schools in which the principal did not employ any criteria for class division. The number of teachers after this filtration was 101,379 and that of students was 437,413 in the 5th grade (Elementary School) and 370,918 in the 9th grade (Middle School).

Second, Statistical Package for the Social Science - SPSS (version 20), defined as a data analysis software, was used for the statistical analysis of PSM, using the Euclidean distance and NNM, to group the classes in pairs. The pairs found were approximately equal when controlling for the chosen variables that affect school performance.

At first, there was an interest in simply calculating the Euclidean distance of all the groups, divided into pairs, and selecting for each of them its pair with the smallest distance between them. As it was not possible to obtain a computational resource for this excessively huge calculation (for example, with 10,000 classes, a distance matrix of 100,000,000 cells would be necessary), we chose to use the k-means clustering method.

Since this study is concerned with binary treatment, it implemented the PSM via a logit model. The logit estimates were treated as forms of approximation and works with the assumption that the data are in not normal distribution.

According to Ref. [40]; correlated binary data appear across models, ranging from those with multivariate measurements, inserted into a random cross-section of arguments, to those that make repeated measurements over time. They also explain how the logit and probit models allows the correlation structure to be modelled flexibly, and the parameters to be interpreted.

It is important to mention that we tested the probit model too, and the result was similar to that of the logit model. However, we opted to use the logit model because of its more adequate precision for this type of study [41]. claim that the probit model use random variation and treatment of factors correspondents over time, being applied in studies that vary in time. They also demonstrate how the logit model is important for studies that is no variation along time and allows overcoming correlated errors. Thus, the balancing test developed in this study was performed in a manner similar to the uninterrupted case, with the imposition of constraint being the only difference, such that the test was performed with mutual support.

After the balancing test was conducted, and the propensity score was estimated, the average treatment was estimated with the treated data, by selecting each treated unit, and identifying a control unit with a similar propensity score. After this, the average of the difference between the response variables of the treated data and matches of the treated observations made over time was obtained. In this sense, the pre-treatment variables, that is, the intervening variables that impact the treatment being developed, used for the combination of classes with similar influence on their scores, presented a significance level of 0.000, that is, both presented the same significance within the study.

To compare clusters, it should be noted that in this framework of choices, it was necessary to arbitrate an initial cluster number. It was decided that N/2 would be appropriate, based on the hypothesis that we would have N/2 clusters with 2 cases each, that is, perfect pairs in the confounding variables. It was defined as N/2 because it is a binary comparison of classes, the classes were compared in

pairs, according to the method of [39]. In a cluster of n = 5, 10 pairs of comparisons of classes t were performed. (t1-t2; t1-t3; t1-t4; t1-t5; t2-t3; t2-t4; t2-t5; t3-t4; t3-t5 and t4-t5).

Thus, for the 17,533 classes in the 5th grade (elementary school), a construction of 8,766 clusters was arbitrated, and for the 13,608 classes in the 9th grade (middle school), the construction of 6,804 clusters was arbitrated.

7. Results

The analysis of the results began with the exclusion of classes that could not be grouped with another class, that is, they formed a cluster with only 1 case. By contrast, although most clusters had at least 2 cases (as desired), approximately 15% of the clusters had more than 2 cases; that is, most of them had 3 or 4 cases. In this instance, as many clusters of pairs as possible were created at the intersection of classes (for example, a cluster of 4 classes formed 6 pairs of classes – the result of combining 4 cases, 2 to 2 – C4.2; Table 2).

Thus, after this construction, 20,362 and 13,038 pairs (or pairs of classes) were created for the 5th and 9th grades, respectively. Once the classes were paired, it was necessary to select a comparative hypothesis test for the paired samples. To that end, it was also necessary to verify the condition of adherence to the normal distribution of the data to define whether they entailed a parametric or non-parametric analysis. In this study, we used the Kolmogorov–Smirnov test to do so and the null hypothesis of normality was rejected. Therefore, a non-parametric analysis was used as illustrated in Table 3.

[42] explains that there are two Kolmogorov Smirnov tests, and the first one serves to test whether a variable follows a given distribution in a population. The author further states that this distribution is usually, but not always, a normal distribution. Therefore, the first test serves to identify whether or not the distribution is normal. The second test is to test whether a variable has an identical distribution between two different samples. Both tests, according to the author, can be used in SPSS. In this study, the former was used to identify the distribution in a population.

The scores for Portuguese language and mathematics in the statistic column of Table 3 have a value greater than 0.01, indicating the rejection of the null hypothesis of adherence to the normal curve, giving rise to a non-parametric analysis. Thus, the non-parametric Wilcoxon test for 2 paired samples was selected.

Overall, it was observed that class size does influence academic performance. When this influence exists, larger classes tend to perform better. Table 4 depicts the results observed for the Portuguese language course in the early grades.

In the Portuguese language of the early grades, it was found that when comparing classes of up to and more than 16 students per class, larger classes performed better. This finding was consistent with the comparison of classes with up to and more than 17 and 20 students per class. However, when comparing classes up to and more than 21 students per class, smaller classes performed better, making the result more robust. This finding was consistent with the comparisons of classes of up to and more than 22, 23, 34, and 35 students per class. These findings suggest that the ideal class size for the Portuguese language in the early grade is 21. Table 5 depicts the results observed for mathematics in the initial grades.

In mathematics, for the initial grades, it was found that when comparing classes of up to and more than 16 students per class, larger classes performed better. However, when comparing classes of up to and more than 29, 30, 31, 32, 33, 34, 35, 36, and 43 students, smaller classes performed better. These findings suggest that the class size for mathematics for the initial grades must be greater than 17 and less than or equal to 29. The best median found in the model (223.6133) corresponds to classes with up to 29 students. Table 6 depicts the results observed during the final grades of the Portuguese language.

Analysis of the Portuguese language results for the final grades found a significant difference in performance in classes of up to and more than 15 students per class. In this case, there is a difference between the medians of 3.0929, which represents a significant difference in performance between smaller and larger classrooms. This indicates that larger classrooms perform better. It was also observed that when the class was larger than 15 students, the performance remained constant, without statistically significant differences up to a class with 27 students. It was also observed that in comparisons of classes of up to and more than 28, 29, 30, and 31 students per class, larger classes perform better. For the range of 32 to 38 students per class, it was possible to maintain performance without statistically significant changes. In the comparison for the range of up to 39–41 students per class, larger classes performed better. Performance was again stable when comparing classes of up to 42 and 43 students per class. In classes with up to 44 and 45 students, larger classrooms performed better. For the range of up to 46–49 students per class, the performance remained statistically stable. The highest median (261.8486) was found in classes with more than 44 students. The findings suggest that the class size for the Portuguese language in the final grades must be 45 students. The results for mathematics in the final grades are presented in Table 7.

Table 2
Clustering data - example.

Cluster 1 = 4 Classes			
Class 1	Comparison		
Class 2	Class 1	vs	Class 2
Class 3	Class 1	vs	Class 3
Class 4	Class 1	vs	Class 4
	Class 2	vs	Class 3
	Class 2	vs	Class 4
	Class 3	vs	Class 4
	Total Comparison = 6		

Table 3
Kolmogorov Smirnov Normality Test for 5th and 9th grade data.

	Statistic	p-value
5th grade		
Portuguese language results	0,010	<0,01
Math results	0,018	<0,01
9th grade		
Portuguese language results	0,031	<0,01
Math results	0,015	<0,01

Table 4
Comparison between class sizes on Portuguese Language 5th grade.

Portuguese Language 5th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 16 students	881	210,5574	−3,32	0,012 ($p < 0,05$)
Above 16 students		213,8777		
Up to 17 students	1.210	210,5574	−2,43	0,017 ($p < 0,05$)
Above 17 students		212,9862		
Up to 20 students	2.318	210,5574	−0,67	0,023 ($p < 0,05$)
Above 20 students		211,2237		
Up to 21 students	2.798	211,4544	0,25	0,009 ($p < 0,05$)
Above 21 students		211,2091		
Up to 22 students	3.288	212,0640	0,50	0,040 ($p < 0,05$)
Above 22 students		211,5619		
Up to 23 students	3.736	212,4137	0,49	0,036 ($p < 0,05$)
Above 23 students		211,9233		
Up to 34 students	1.740	213,0719	2,37	0,002 ($p < 0,05$)
Above 34 students		210,6946		
Up to 35 students	1.221	212,9300	4,10	0,002 ($p < 0,05$)
Above 35 students		208,8341		

Table 5
Comparison between class sizes on Math 5th grade.

Math 5th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 16 students	881	220,1121	−2,3398	0,021 ($p < 0,05$)
Above 16 students		222,4519		
Up to 17 students	1.210	218,6385	−3,3244	0,033 ($p < 0,05$)
Above 17 students		221,9629		
Up to 29 students	4.482	223,6133	1,7013	0,017 ($p < 0,05$)
Above 29 students		221,912		
Up to 30 students	4.068	222,9976	1,0856	0,031 ($p < 0,05$)
Above 30 students		221,912		
Up to 31 students	3.562	221,8295	1,1659	0,002 ($p < 0,05$)
Above 31 students		220,6636		
Up to 32 students	2.828	221,5416	0,891	0,000 ($p < 0,05$)
Above 32 students		220,6506		
Up to 33 students	2.328	220,7266	1,3429	0,000 ($p < 0,05$)
Above 33 students		219,3837		
Up to 34 students	1.740	221,104	5,6629	0,000 ($p < 0,05$)
Above 34 students		215,4411		
Up to 35 students	1.221	220,8753	7,1905	0,000 ($p < 0,05$)
Above 35 students		213,6848		
Up to 36 students	760	218,3807	6,9254	0,000 ($p < 0,05$)
Above 36 students		211,4553		
Up to 43 students	29	215,7309	8,8435	0,039 ($p < 0,05$)
Above 43 students		206,8874		

Analysis of the mathematics subject results of the final grades indicates that larger classes performed better. However, the performance of smaller classes with up to 32 students remained statistically stable, while classes in the range of 33–40 students per class experienced reduced performance. Thus, classes with up to 32 students performed better than those with the range of 33 to 40 students. For classes with 41 students or more, it was observed that larger classes performed better. Thus, in order to find the ideal number of

Table 6
Comparison between class sizes on Portuguese Language 9th grade.

Portuguese Language 9th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 15 students	400	246,3518	−3,0929	0,020 ($p < 0,05$)
Above 15 students		249,4447		
Up to 28 students	2.816	255,4747	−0,96	0,017 ($p < 0,05$)
Above 28 students		256,4335		
Up to 29 students	2.948	254,6743	−2,3237	0,000 ($p < 0,05$)
Above 29 students		256,9980		
Up to 30 students	2.927	254,6743	−2,49	0,000 ($p < 0,05$)
Above 30 students		257,1623		
Up to 31 students	2.884	255,6780	−1,84	0,002 ($p < 0,05$)
Above 31 students		257,5164		
Up to 39 students	780	252,0462	−3,6615	0,007 ($p < 0,05$)
Above 39 students		255,7077		
Up to 40 students	604	252,2976	−4,8647	0,010 ($p < 0,05$)
Above 40 students		257,1623		
Up to 41 students	327	251,2653	−6,3632	0,047 ($p < 0,05$)
Above 41 students		257,6285		
Up to 44 students	116	252,7799	−9,0687	0,000 ($p < 0,05$)
Above 44 students		261,8486		
Up to 45 students	72	251,3027	−8,1588	0,003 ($p < 0,05$)
Above 45 students		259,4615		

Table 7
Comparison between class sizes on Math 9th grade.

Math 9th grade				
Logic	N	Median	Median difference (up to – above)	Median difference (up to – above)
Up to 32 students	2.870	253,8103	0,04	0,04
Above 32 students		253,7727		
Up to 33 students	2.732	254,5607	0,7880	0,002 ($p < 0,05$)
Above 33 students		253,7727		
Up to 34 students	2.483	253,8576	0,6578	0,012 ($p < 0,05$)
Above 34 students		253,1998		
Up to 36 students	1.808	253,4177	0,3099	0,038 ($p < 0,05$)
Above 36 students		253,1078		
Up to 37 students	1.356	252,2840	4,1928	0,001 ($p < 0,05$)
Above 37 students		248,0912		
Up to 38 students	1.045	250,5095	2,2620	0,017 ($p < 0,05$)
Above 38 students		248,2475		
Up to 41 students	327	247,2052	−7,2269	0,034 ($p < 0,05$)
Above 41 students		254,4321		
Up to 44 students	116	247,1770	−20,35	0,000 ($p < 0,05$)
Above 44 students		267,5242		
Up to 45 students	72	243,5252	−21,62	0,003 ($p < 0,05$)
Above 45 students		265,1358		
Up to 47 students	29	244,0724	−29,1185	0,028 ($p < 0,05$)
Above 47 students		273,1909		

students per class, we chose to find the highest median, which was 273.19 being watched in the comparison of classes with up to or more than 47 students per class. Therefore, the ideal class size for mathematics in the final grades must be greater than 47 students—a number that presents the maximum value of the medians.

8. Discussion

There are several studies that confirm the influence of the number of students per class on school performance is significant, but usually these have a relatively small amount of data. This is the case of studies such as the one by Ref. [23] in Hong Kong, which used a sample of only 37 schools, the one by Ref. [24] who used 69 schools and reported that class size influenced performance only in the first year of study and over time this influence dissipated and [22] from the Netherlands who, using a sample of 416 schools, began to observe little influence of smaller classes on performance [22–24].

When data is relatively large, as in the case of Bangladesh with 12,614 schools, the evidence on the influence of the number of students per class was not significant. Similar results were also observed in Ukraine's 11,683 schools. Both studies used Ordinary Least Squares (OLS) model and found no significant evidence supporting the proposal that minor classes perform better. In this vein, an

Africa study with a total 8,626 schools, used the OLS model, but did not find evidence to claim that the class size influences achievement [43, 44, 45].

A comprehensive analysis carried out by Ref. [46] with evidence from the Progress in International Reading Literacy Study (PIRLS) for the period 2001 to 2011. This analysis based on 8 European countries (Bulgaria, Germany, Hungary, Italy, Lithuania, Romania, Slovakia, and Slovenia) found that there were no statistically significant results in 7 of these countries regarding the influence of class size on school performance, even when distortions were verified over the years. The only country where this evidence was found was Romania, which showed that smaller classes increased reading performance [46].

[47] also demonstrated this by investigating the results of 277 studies on the effects of the number of students per class on school performance. He found that it is possible to verify the effect when the studies presented a statistically significant p-value, 14% of the studies had an influence on performance in terms of class size, another 14% had no influence, and in 72% of the cases, it was possible to notice that the studies did not even reach any conclusion. This demonstrates the challenge of finding a single answer in terms of the influence of the number of students per class on school performance that would fit in any teaching situation. Along the same lines, [48] investigated 47 countries using data from the Trends in International Mathematics and Science Study - TIMSS to verify if there was a relationship between class size and performance and found statistically significant hypotheses in only 14 of the verified countries, and they also found that the effect of class size is smaller in countries with higher teacher quality [47, 48].

9. Conclusion

Even though, prior studies have not found a direct relationship between class size and school performance, we were still interested in testing whether this hypothesis would apply to Brazil. In addition, with the amount of data used in this study, it is of paramount importance for literature to demonstrate the behavior of this hypothesis in the Brazilian context.

This study's findings agree with that of [22]; and demonstrate that students in larger classes perform better than those in smaller classes. This study's findings, in tandem with that of Bosker's, has the advantage of avoiding the Hawthorne effect. According to Woesmann and West (2002), the Hawthorne effect is characterized by participants increasing their efforts when they know they are being studied, which would not happen under normal circumstances. Both surveys analyzed microdata with random selection criteria and without the students knowing that they had participated in the survey, and both showed the same result. These results contradict what common sense tells us.

It should be noted that, initially, the interest was in calculating the Euclidean distance between groups, by dividing them into pairs and, selecting the pair with minimum distance between themselves. However, this was not possible, as the computational limitations of the Propensity Score Matching estimation algorithm in the SPSS software led us to use the k-means grouping method to avoid limiting complete comparison between all classes. If this was not adopted, 10,000 classes would require a distance matrix of 100,000,000 cells.

This limitation is similar to that of Monte Carlo's study (2010), wherein the researchers used a PC (CPU: Pentium Dual-Core E5400; memory: 4 GB). Due to computational limitations, approximately 10 days were spent in just forming combined samples of Propensity Score Matching scores [49]. After 12 years since Monte Carlo's study, even with the use of a PC (CPU: Pentium 12-Core i7, memory: 16 GB), it remained impossible to make such a large comparison. However, considering the accelerated pace of technological evolution, one can believe that it shall be possible to make this comparison in the near future.

[13] opines that in situations with specific teachers, groups of students, and subjects, it is likely that smaller classes perform better. Thus, this research contributes to educational literature by demonstrating that the class size does not influence the school performance of Brazilian Elementary and Middle school students, and to class policies to improve the scope of education. Thus like [13]; we believe that if a different method of teaching were to be used in Brazil, the number of students per classroom could become an important determinant of performance.

In conclusion, this manuscript contributes significantly to clarifying whether the mere decrease or increase in the number of students per class can improve educational quality from the perspective of increased performance. The findings demonstrate that this measure alone is not enough to achieve this school improvement.

Author contribution statement

André Calixto Gonçalves: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Leandro Campi Prearo: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data.

Nonato Assis de Miranda; Paulo Roberto Mendes: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

The database is available on the brazilian government website, it is microdata referring to large-scale assessments carried out in

brazil for all years. If requested, we can send them. But everyone has access to them

Declaration of interest's statement

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e15130.

Appendix

Table A1
Comparison between class sizes of Portuguese Language in the 5th grade

Portuguese Language 5th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 08 students	5	224,77	0,01	0,686 (p > 0,05)
Above 08 students		224,76		
Up to 09 students	12	235,65	9,67	0,209 (p > 0,05)
Above 09 students		225,98		
Up to 10 students	13	235,65	10,89	0,133 (p > 0,05)
Above 10 students		224,76		
Up to 11 students	22	224,77	14,19	0,758 (p > 0,05)
Above 11 students		210,58		
Up to 12 students	102	205,09	−4,99	0,932 (p > 0,05)
Above 12 students		210,08		
Up to 13 students	211	211,08	0,99	0,730 (p > 0,05)
Above 13 students		210,09		
Up to 14 students	369	211,41	1,55	0,071 (p > 0,05)
Above 14 students		209,86		
Up to 15 students	588	210,56	−1,71	0,403 (p > 0,05)
Above 15 students		212,27		
Up to 16 students	**881**	**210,5574**	**−3,32**	**0,012 (p < 0,05)**
Above 16 students		**213,8777**		
Up to 17 students	**1.210**	**210,5574**	**−2,43**	**0,017 (p < 0,05)**
Above 17 students		**212,9862**		
Up to 18 students	1.540	210,55	−0,85	0,275 (p > 0,05)
Above 18 students		211,4		
Up to 19 students	1.923	210,35	−0,33	0,071 (p > 0,05)
Above 19 students		210,68		
Up to 20 students	**2.318**	**210,5574**	**−0,67**	**0,023 (p < 0,05)**
Above 20 students		**211,2237**		
Up to 21 students	**2.798**	**211,4544**	**0,25**	**0,009 (p < 0,05)**
Above 21 students		**211,2091**		
Up to 22 students	**3.288**	**212,064**	**0,5**	**0,040 (p < 0,05)**
Above 22 students		**211,5619**		
Up to 23 students	**3.736**	**212,4137**	**0,49**	**0,036 (p < 0,05)**
Above 23 students		**211,9233**		
Up to 24 students	4.182	213,69	1,13	0,416 (p > 0,05)
Above 24 students		212,56		
Up to 25 students	4.435	212,05	−0,65	0,560 (p > 0,05)
Above 25 students		212,7		
Up to 26 students	4.769	214,04	0,72	0,223 (p > 0,05)
Above 26 students		213,32		
Up to 27 students	4.807	214,94	0,86	0,665 (p > 0,05)
Above 27 students		214,08		
Up to 28 students	4.658	215,68	0,96	0,879 (p > 0,05)
Above 28 students		214,72		
Up to 29 students	4.482	215,38	1,29	0,270 (p > 0,05)
Above 29 students		214,09		
Up to 30 students	4.068	214,66	0,22	0,352 (p > 0,05)
Above 30 students		214,44		
Up to 31 students	3.562	213,7	0,41	0,294 (p > 0,05)
Above 31 students		213,29		
Up to 32 students	2.828	213,73	−0,16	0,189 (p > 0,05)
Above 32 students		213,89		

(*continued on next page*)

Table A1 (*continued*)

Portuguese Language 5th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 33 students	2.328	212,83	0,23	0,146 (p > 0,05)
Above 33 students		212,6		
Up to 34 students	**1.740**	**213,0719**	**2,37**	**0,002 (p < 0,05)**
Above 34 students		**210,6946**		
Up to 35 students	**1.221**	**212,93**	**4,1**	**0,002 (p < 0,05)**
Above 35 students		**208,8341**		
Up to 36 students	760	211,56	2,99	0,150 (p > 0,05)
Above 36 students		208,57		
Up to 37 students	441	208,04	3,12	0,985 (p > 0,05)
Above 37 students		204,92		
Up to 38 students	293	208,02	−0,09	0,223 (p > 0,05)
Above 38 students		208,11		
Up to 39 students	204	205,99	−1,6	0,365 (p > 0,05)
Above 39 students		207,59		
Up to 40 students	127	208,93	4,01	0,364 (p > 0,05)
Above 40 students		204,92		
Up to 41 students	69	205,77	8,25	0,749 (p > 0,05)
Above 41 students		197,52		
Up to 42 students	57	205,77	13,27	0,682 (p > 0,05)
Above 42 students		192,5		
Up to 43 students	29	209,06	−6,97	0,456 (p > 0,05)
Above 43 students		216,03		
Up to 44 students	23	207,34	−8,69	0,927 (p > 0,05)
Above 44 students		216,03		
Up to 45 students	18	204,92	−17,67	0,085 (p > 0,05)
Above 45 students		222,59		

Table A2
Comparison between class sizes of Math in the 5th grade

Math 5th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 08 students	5	223,8	6,13	0,686 (p > 0,05)
Above 08 students		217,67		
Up to 09 students	12	225,62	−11,4	0,158 (p > 0,05)
Above 09 students		237,02		
Up to 10 students	13	225,62	−9,54	0,249 (p > 0,05)
Above 10 students		235,16		
Up to 11 students	22	223,8	5,46	0,661 (p > 0,05)
Above 11 students		218,34		
Up to 12 students	102	210,27	−7,1	0,379 (p > 0,05)
Above 12 students		217,37		
Up to 13 students	211	214,64	−5,24	0,158 (p > 0,05)
Above 13 students		219,88		
Up to 14 students	369	219,9	1,59	0,410 (p > 0,05)
Above 14 students		218,31		
Up to 15 students	588	217,56	−2,73	0,056 (p > 0,05)
Above 15 students		220,29		
Up to 16 students	**881**	**220,1121**	**−2,3398**	**0,021 (p < 0,05)**
Above 16 students		**222,4519**		
Up to 17 students	**1.210**	**218,6385**	**−3,3244**	**0,033 (p < 0,05)**
Above 17 students		**221,9629**		
Up to 18 students	1.540	219,14	−0,76	0,693 (p > 0,05)
Above 18 students		219,9		
Up to 19 students	1.923	218,98	−0,13	0,438 (p > 0,05)
Above 19 students		219,11		
Up to 20 students	2.318	219,58	0,06	0,278 (p > 0,05)
Above 20 students		219,52		
Up to 21 students	2.798	221,25	1,73	0,117 (p > 0,05)
Above 21 students		219,52		
Up to 22 students	3.288	222,19	2,15	0,720 (p > 0,05)
Above 22 students		220,04		
Up to 23 students	3.736	222,19	1,68	0,894 (p > 0,05)
Above 23 students		220,51		
Up to 24 students	4.182	222,89	2,04	0,151 (p > 0,05)
Above 24 students		220,85		
Up to 25 students	4.435	222,45	−0,06	0,062 (p > 0,05)

(*continued on next page*)

Table A2 (*continued*)

Math 5th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Above 25 students		222,51		
Up to 26 students	4.769	223,41	1,83	0,405 (p > 0,05)
Above 26 students		221,58		
Up to 27 students	4.807	223,62	1,63	0,174 (p > 0,05)
Above 27 students		221,99		
Up to 28 students	4.658	224,07	1,66	0,179 (p > 0,05)
Above 28 students		222,41		
Up to 29 students	**4.482**	**223,6133**	**1,7013**	**0,017 (p < 0,05)**
Above 29 students		**221,912**		
Up to 30 students	**4.068**	**222,9976**	**1,0856**	**0,031 (p < 0,05)**
Above 30 students		**221,912**		
Up to 31 students	**3.562**	**221,8295**	**1,1659**	**0,002 (p < 0,05)**
Above 31 students		**220,6636**		
Up to 32 students	**2.828**	**221,5416**	**0,891**	**0,000 (p < 0,05)**
Above 32 students		**220,6506**		
Up to 33 students	**2.328**	**220,7266**	**1,3429**	**0,000 (p < 0,05)**
Above 33 students		**219,3837**		
Up to 34 students	**1.740**	**221,104**	**5,6629**	**0,000 (p < 0,05)**
Above 34 students		**215,4411**		
Up to 35 students	**1.221**	**220,8753**	**7,1905**	**0,000 (p < 0,05)**
Above 35 students		**213,6848**		
Up to 36 students	**760**	**218,3807**	**6,9254**	**0,000 (p < 0,05)**
Above 36 students		**211,4553**		
Up to 37 students	441	215,73	8,29	0,052 (p > 0,05)
Above 37 students		207,44		
Up to 38 students	293	215,55	8,67	0,372 (p > 0,05)
Above 38 students		206,88		
Up to 39 students	204	214,63	7,75	0,148 (p > 0,05)
Above 39 students		206,88		
Up to 40 students	127	216,55	9,67	0,373 (p > 0,05)
Above 40 students		206,88		
Up to 41 students	69	214,41	10,21	0,098 (p > 0,05)
Above 41 students		204,2		
Up to 42 students	57	214,41	16,64	0,055 (p > 0,05)
Above 42 students		197,77		
Up to 43 students	**29**	**215,7309**	**8,8435**	**0,039 (p < 0,05)**
Above 43 students		**206,8874**		
Up to 44 students	23	214,41	7,52	0,191 (p > 0,05)
Above 44 students		206,89		
Up to 45 students	18	213,42	6,54	0,647 (p > 0,05)
Above 45 students		206,88		

Table A3
Comparison between class sizes of Portuguese Language in the 9th grade

Portuguese Language 9th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 12 students	97	244,11	−10,14	0,970 (p > 0,05)
Above 12 students		254,25		
Up to 13 students	194	240,01	−7,53	0,063 (p > 0,05)
Above 13 students		247,54		
Up to 14 students	286	245,02	−2,88	0,281 (p > 0,05)
Above 14 students		247,9		
Up to 15 students	**400**	**246,3518**	**−3,0929**	**0,020 (p < 0,05)**
Above 15 students		**249,4447**		
Up to 16 students	475	247,97	−2,66	0,074 (p > 0,05)
Above 16 students		250,63		
Up to 17 students	674	252,67	1,09	0,957 (p > 0,05)
Above 17 students		251,58		
Up to 18 students	840	250,04	−1,53	0,379 (p > 0,05)
Above 18 students		251,57		
Up to 19 students	1.048	253,14	0,79	0,235 (p > 0,05)
Above 19 students		252,35		
Up to 20 students	1.199	252,54	0,07	0,052 (p > 0,05)
Above 20 students		252,47		
Up to 21 students	1.398	253,48	1,73	0,718 (p > 0,05)
Above 21 students		251,75		

(*continued on next page*)

Table A3 (*continued*)

Portuguese Language 9th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 22 students	1.629	255,15	2,64	0,822 (p > 0,05)
Above 22 students		252,51		
Up to 23 students	1.873	253,48	−0,02	0,534 (p > 0,05)
Above 23 students		253,5		
Up to 24 students	2.142	255,69	1,51	0,490 (p > 0,05)
Above 24 students		254,18		
Up to 25 students	2.322	255,69	1,38	0,850 (p > 0,05)
Above 25 students		254,31		
Up to 26 students	2.500	255,69	0,33	0,706 (p > 0,05)
Above 26 students		255,36		
Up to 27 students	2.661	255,69	−0,01	0,351 (p > 0,05)
Above 27 students		255,7		
Up to 28 students	**2.816**	**255,4747**	**−0,96**	**0,017 (p < 0,05)**
Above 28 students		**256,4335**		
Up to 29 students	**2.948**	**254,6743**	**−2,3237**	**0,000 (p < 0,05)**
Above 29 students		**256,998**		
Up to 30 students	**2.927**	**254,6743**	**−2,49**	**0,000 (p < 0,05)**
Above 30 students		**257,1623**		
Up to 31 students	**2.884**	**255,678**	**−1,84**	**0,002 (p < 0,05)**
Above 31 students		**257,5164**		
Up to 32 students	2.870	256,29	−1,22	0,114 (p > 0,05)
Above 32 students		257,51		
Up to 33 students	2.732	256,68	−0,94	0,396 (p > 0,05)
Above 33 students		257,62		
Up to 34 students	2.483	256,31	−1,21	0,359 (p > 0,05)
Above 34 students		257,52		
Up to 35 students	2.207	255,92	−1,24	0,949 (p > 0,05)
Above 35 students		257,16		
Up to 36 students	1.808	255,88	−1,3	0,797 (p > 0,05)
Above 36 students		257,18		
Up to 37 students	1.356	255,19	1,05	0,130 (p > 0,05)
Above 37 students		254,14		
Up to 38 students	1.045	254,08	−0,09	0,540 (p > 0,05)
Above 38 students		254,17		
Up to 39 students	**780**	**252,0462**	**−3,6615**	**0,007 (p < 0,05)**
Above 39 students		**255,7077**		
Up to 40 students	**604**	**252,2976**	**−4,8647**	**0,010 (p < 0,05)**
Above 40 students		**257,1623**		
Up to 41 students	**327**	**251,2653**	**−6,3632**	**0,047 (p < 0,05)**
Above 41 students		**257,6285**		
Up to 42 students	208	251,7	−1,3	0,281 (p > 0,05)
Above 42 students		253		
Up to 43 students	169	254,21	−5,25	0,525 (p > 0,05)
Above 43 students		259,46		
Up to 44 students	**116**	**252,7799**	**−9,0687**	**0,000 (p < 0,05)**
Above 44 students		**261,8486**		
Up to 45 students	**72**	**251,3027**	**−8,1588**	**0,003 (p < 0,05)**
Above 45 students		**259,4615**		
Up to 46 students	35	252,68	−6,78	0,272 (p > 0,05)
Above 46 students		259,46		
Up to 47 students	29	252,87	−6,59	0,294 (p > 0,05)
Above 47 students		259,46		
Up to 48 students	26	252,11	−7,35	0,454 (p > 0,05)
Above 48 students		259,46		
Up to 49 students	19	252,87	−6,59	1,000 (p > 0,05)
Above 49 students		259,46		

Table A4
Comparison Between class sizes of Math in the 9th grade

Math 9th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Up to 12 students	97	250,82	1,04	0,265 (p > 0,05)
Above 12 students		249,78		
Up to 13 students	194	243,23	−2,45	0,126 (p > 0,05)
Above 13 students		245,68		
Up to 14 students	286	243,51	−2,1	0,510 (p > 0,05)

(*continued on next page*)

Table A4 (*continued*)

Math 9th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Above 14 students		245,61		
Up to 15 students	400	243,85	−2,41	0,060 (p > 0,05)
Above 15 students		246,26		
Up to 16 students	475	247,75	−0,1	0,216 (p > 0,05)
Above 16 students		247,85		
Up to 17 students	674	249,17	−0,06	0,505 (p > 0,05)
Above 17 students		249,23		
Up to 18 students	840	247,01	−2,32	0,858 (p > 0,05)
Above 18 students		249,33		
Up to 19 students	1.048	249,1	−0,71	0,498 (p > 0,05)
Above 19 students		249,81		
Up to 20 students	1.199	249,11	−0,86	0,772 (p > 0,05)
Above 20 students		249,97		
Up to 21 students	1.398	249,78	0,89	0,187 (p > 0,05)
Above 21 students		248,89		
Up to 22 students	1.629	250,51	1,06	0,685 (p > 0,05)
Above 22 students		249,45		
Up to 23 students	1.873	251,07	1,24	0,537 (p > 0,05)
Above 23 students		249,83		
Up to 24 students	2.142	252,66	1,46	0,223 (p > 0,05)
Above 24 students		251,2		
Up to 25 students	2.322	252,68	1,04	0,172 (p > 0,05)
Above 25 students		251,64		
Up to 26 students	2.500	252,76	0,69	0,527 (p > 0,05)
Above 26 students		252,07		
Up to 27 students	2.661	252,76	0,62	0,706 (p > 0,05)
Above 27 students		252,14		
Up to 28 students	2.816	252,76	−0,43	0,753 (p > 0,05)
Above 28 students		253,19		
Up to 29 students	2.948	252,6	−0,94	0,248 (p > 0,05)
Above 29 students		253,54		
Up to 30 students	2.927	252,76	−0,78	0,862 (p > 0,05)
Above 30 students		253,54		
Up to 31 students	2.884	253,27	−0,79	0,715 (p > 0,05)
Above 31 students		254,06		
Up to 32 students	**2.870**	**253,8103**	**0,04**	**0,013 (p < 0,05)**
Above 32 students		**253,7727**		
Up to 33 students	**2.732**	**254,5607**	**0,788**	**0,002 (p < 0,05)**
Above 33 students		**253,7727**		
Up to 34 students	**2.483**	**253,8576**	**0,6578**	**0,012 (p < 0,05)**
Above 34 students		**253,1998**		
Up to 35 students	2.207	253,36	0,16	0,107 (p > 0,05)
Above 35 students		253,2		
Up to 36 students	**1.808**	**253,4177**	**0,3099**	**0,038 (p < 0,05)**
Above 36 students		**253,1078**		
Up to 37 students	**1.356**	**252,284**	**4,1928**	**0,001 (p < 0,05)**
Above 37 students		**248,0912**		
Up to 38 students	**1.045**	**250,5095**	**2,262**	**0,017 (p < 0,05)**
Above 38 students		**248,2475**		
Up to 39 students	780	249,57	−2,68	0,472 (p > 0,05)
Above 39 students		252,25		
Up to 40 students	604	248,71	−5,4	0,141 (p > 0,05)
Above 40 students		254,11		
Up to 41 students	**327**	**247,2052**	**−7,2269**	**0,034 (p < 0,05)**
Above 41 students		**254,4321**		
Up to 42 students	208	247,71	−0,52	0,315 (p > 0,05)
Above 42 students		248,23		
Up to 43 students	169	249,89	−4,54	0,475 (p > 0,05)
Above 43 students		254,43		
Up to 44 students	**116**	**247,177**	**−20,35**	**0,000 (p < 0,05)**
Above 44 students		**267,5242**		
Up to 45 students	**72**	**243,5252**	**−21,62**	**0,003 (p < 0,05)**
Above 45 students		**265,1358**		
Up to 46 students	35	242,97	−4,21	0,55 (p > 0,05)
Above 46 students		247,18		
Up to 47 students	**29**	**244,0724**	**−29,1185**	**0,028 (p < 0,05)**
Above 47 students		**273,1909**		
Up to 48 students	26	242,6	−21,48	0,58 (p > 0,05)

(*continued on next page*)

Table A4 (*continued*)

Math 9th grade				
Logic	N	Median	Median difference (up to – above)	p-value (Wilcoxon test)
Above 48 students		264,08		
Up to 49 students	19	244,07	−29,12	0,147 (p > 0,05)
Above 49 students		273,19		

References

[1] E. Word, The State of Tennessee's Student/Teacher Achievement Ratio – STAR Project, Technical Report, Tennessee, 1990, pp. 1985–1990. Available from: https://eric.ed.gov/?id=ED328356.

[2] G.W. Bohrnstedt, B.M. Stecher, Class Size Reduction C California: Findings from 1999–00 and 2000–01, vol. 1, CSR Research Consortium, California, USA, 2002, p. 160. Available from: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.694.7009&rep=rep1&type=pdf. (Accessed 15 November 2021).

[3] Peter Maier, Alex Molnar, Percy Stephen, Phillip Smith, John Zahorik, First Year Results Of The Student Achievement Guarantee In Education Program, SAGE Evaluation Team Wisconsin, Dez, 1997. Available from: https://eric.ed.gov/?id=ED441002.

[4] David A. Gilman, Prime time in the first grade at the north gibson school corporation: the first four years, Resour. Edu. 25 (1988), 1 USA, jan.

[5] Nina Bascia, Reducing Class Size: what Do We Know? *Canadian Education Association*, Canada, 2010, p. 22. Available from: https://eric.ed.gov/?id=ED573360.

[6] Sahin B. Ece, Dostoglu Neslihan, Evaluation of kindergarten group rooms in the context of size: children and teacher's perspective in Turkey, Eur. J. Contemp. Educ. 10 (4) (2014) 253–264, https://doi.org/10.13187/ejced.2014.10.253. Turkey.

[7] p. Fredrikson, b. ockert, h. oosterbeek, Long-Terms Effects of Class Size, IZA Discussion Paper, Sweden, jul.2011, p. 5879. Available from: https://ideas.repec.org/p/iza/izadps/dp5879.html.

[8] Sule, S. Samuel, Effects of assignment and class size on secondary school student's achievement in mathematics, Nigeria, ATBU J. Sci. Tech. Educ. 4 (2) (jun.2016). Available from: https://www.atbuftejoste.com/index.php/joste/article/view/189.

[9] Victor Oluwatosin Ajayi, Christina Audu, Evelyn Ajayi, Influence of class size on students' classroom discipline, engagement and communication: a case study of senior secondary schools in ekiti state, Nigeria, Sky J. Educ. Res. 5 (5) (November 7, 2017) 34–41, 11/2017, Available at:: SSRN: https://ssrn.com/abstract=3066387.

[10] Carmel Blank, Reut Shafrir, Yossi Shavit, Is less really more? On the relationship between class size and educational achievement in Israel, TAUB Center for Social Studies in Israel Israel, 2016. Available from: https://www.taubcenter.org.il/en/research/is-less-really-more-on-the-relationship-between-class-size-and-educational-achievement-in-israel/.

[11] Peter Blatchford, Paul Basset, Brown Penelope, Examining the Effect of Class Size on Classroom Engagement and Teacherepupil Interaction: Differences in Relation to Pupil Prior Attainment and Primary vs. Secondary Schools, vol. 21 n.6, *ELSEVIER*, London, dez, 2010, https://doi.org/10.1016/j.learninstruc.2011.04.001.

[12] Yusuf, A. Taofeek, C.A. Onifade, O.S. Bello, Impact of class size on learning, behavioral and general attitudes of students in secondary schools in abeokuta, Nigeria, J. Res. Init. 2 (1) (2016) 19. Available from: https://www.researchgate.net/publication/298773857_Impact_of_Class_Size_on_Learning_Behavioral_and_General_Attitudes_of_Students_in_Secondary_Schools_in_Abeokuta_Ogun_State_Nigeria.

[13] E.A. Hanushek, The Evidence on Class Size, Allen Wallis Institute of Political Economy, University of Rochester, Occasional Paper Number 98-1, 1998. Available from: https://econpapers.repec.org/paper/rocwallis/wp10.htm.

[14] Brazil, Lei nº 9.394, de 20 de dezembro de 1996. Estabelece diretrizes e bases da educação nacional, 1996. Available from: http://www.planalto.gov.br/ccivil_03/leis/l9394.htm. (Accessed 6 December 2021). Brasília, DF: Diário Oficial da União, 23 Dezember.

[15] Brazil Lei nº 13257, de 08 de março de 2016. Dispõe sobre as políticas públicas para a Primeira Infância. Diário Oficial da União, Brasília, DF, 8 Mar. Available from: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/lei/l13257.htm Accessed December 6, 2021.

[16] Brazil, Diretrizes curriculares nacionais da educação basica, Ministério da Educação, Secretaria de Educação Básica, Brasília, 2013, p. 542. Available from: http://portal.mec.gov.br/docman/julho-2013-pdf/13677-diretrizes-educacao-basica-2013-pdf/file. (Accessed 6 December 2021).

[17] Brazil, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Censo da educação básica 2020. Notas estatísticas, 2020. Available from: https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_escolar_2020.pdf. (Accessed 6 December 2021) (Brasília, DF: INEP).

[18] OECD, Education Policy Outlook: Brazil—With a Focus on National and Subnational Policies, 2021. Available from: https://www.oecd.org/education/policy-outlook/country-profile-Brazil-2021-INT-PT.pdf. (Accessed 12 December 2021).

[19] Brazil, Brazil. Instituto Nacional De Estudos E Pesquisas Educacionais Anísio Teixeira (INEP). Prova Brazil, 2017, MEC, Brasília, 2017, 2017. Available from: http://portal.mec.gov.br/prova-brasil. (Accessed 6 December 2021).

[20] G.J. Whitehurst, M.M. Chingos, Class Size. What Research Says and what it Means for State Policy, vol. 1, Brown Center on Education Policy at Brookings, 2011, pp. 1–14. Available from: https://www.brookings.edu/research/class-size-what-research-says-and-what-it-means-for-state-policy/.

[21] E. Graue, K. Hatch, K. Rao, D. One, The wisdom of class-size reduction, Am. Educ. Res. J. 44 (2007) 670–700, https://doi.org/10.3102/0002831207306755.

[22] R.J. Bosker, The class size question in primary schools: policy issues, theory, and empirical findings from The Netherlands, Int. J. Educ. Res. 29 (1998) 763–778, https://doi.org/10.1016/S0883-0355(98)00062-7.

[23] M. Galton, T. Pell, Do class size reductions make a difference to classroom practice? The case of Hong Kong primary schools, Int. J. Educ. Res. 53 (2012) 22–31, https://doi.org/10.1016/j.ijer.2011.12.004.

[24] P. Bressoux, L. Lima, C. Monseur, Reducing the number of pupils in French first-grade classes: is there evidence of contemporaneous and carryover effects? Int. J. Educ. Res. 96 (2019) 136–145, https://doi.org/10.1016/j.ijer.2018.10.006.

[25] A. Ewert, J. Sibthorp, Creating outcomes through experiential education: the challenge of confounding variables, J. Exp. Educ. 31 (2009) 376–389, 10.1177%2F105382590803100305.

[26] Eric A. Hanushek, The Economic Value of Improved Schools, EUA abr, NIET -National Institute for Excellence in Teaching, California, 2019, pp. 59–72. Available from: http://hanushek.stanford.edu/publications/economic-value-improving-public-schools.

[27] Ikaro T. Pass, Dissertação, Evaluation of the Quality Determinants of Basic Education in 2013: a Comparison of the South-Center and West Mesoregions from Paraná, vol. 95, (Mestrado em Desenvolvimento Econômico Regional e Economia do Agronegócio) - Universidade Estadual do Oeste do Paraná, Toledo, 2016, 2016.

[28] J.F. Soares, O efeito da escola no desempenho cognitivo de seus alunos, REICE. Revista Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación 2 (2) (2016). Available from: https://revistas.uam.es/reice/article/view/5550.

[29] Caroline M. Hoxby, The effects of class size and composition on student achievement: new evidence from natural population variation, Inglaterra, Q. J. Econ. 1 (1) (1998). Available from: https://ideas.repec.org/p/nbr/nberwo/6869.html.

[30] Matthew Chingos, The impact of a universal class-size reduction policy: evidence from Florida's statewide mandate, Econ. Educ. Rev. 31 (5) (2012) 543–562, https://doi.org/10.1016/j.econedurev.2012.03.002. EUA, out.

[31] K.F. Punch, A. Oancea, Introduction to Research Methods in Education, Sage Publications, Thousand Oaks, California, 2014. Available from: https://us.sagepub.com/en-us/nam/introduction-to-research-methods-in-education/book239756.
[32] F. Thoemmes, Propensity Score Matching in SPSS, vol. 30, Arxiv P., Germany, jan.2012, https://doi.org/10.48550/arXiv.1201.6385.
[33] S. Guha, N. Mishra, Clustering data streams, in: M. Garofalakis, J. Gehrke, R. Rastogi (Eds.), Data Stream Management. Data-Centric Systems and Applications, Springer, Berlin, Heidelberg, 2016, https://doi.org/10.1007/978-3-540-28608-0_8.
[34] P. Berkhin, A survey of clustering data mining techniques, in: J. Kogan, C. Nicholas, M. Teboulle (Eds.), Grouping Multidimensional Data, Springer, 2006, pp. 25–71, https://doi.org/10.1007/3-540-28349-8_2.
[35] H. Steinhaus, Sur la division des corps matériels en parties, Bulletin de l'Académie Polonaise des Sciences (1956) 801–804. Classe 3, 12.
[36] T. Sorenson, A method of establishing groups of equal amplitudes in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons, K. Dan. Vidensk. Selsk. Biol. Skr. 5 (1948) 1–34.
[37] P.E. Black, Euclidean distance [online], in: P.E. Black (Ed.), Dictionary of Algorithms and Data Structures, 2004. Available from: https://www.nist.gov/dads/HTML/euclidndstnc.html. (Accessed 17 December 2004).
[38] A. Huang, Similarity Measures for Text Document Clustering, *CiteSeerX*, Nova Zelândia, 2008. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.332.4480.
[39] W.G. Cochran, D.B. Rubin, Controlling bias in observational studies: a review, Sankhya 35 (1973) 417–446.
[40] Chib Siddhartha, Edward Greenberg, Analysis of multivariate probit models, Biometrika 85 (2) (1998) 347–361.
[41] Pandit B. Dash, Bighnaraj Naik, Janmenjoy Nayak, S. Vimal, Socio-economic factor analysis for sustainable and smart precision agriculture: an ensemble learning approach, Comput. Commun. 182 (2022) 72–87, https://doi.org/10.1016/j.comcom.2021.11.002. ISSN 0140–3664.
[42] R.G. Van den Berg, SPSS Kolmogorov–Smirnov Test for Normality – the Ultimate Guide, SPSS tutorials, Holanda, 2019. Available from: https://www.spss-tutorials.com/spss-kolmogorov-smirnov-test-for-normality/.
[43] M.N. Asadullah, the effect of class size on student achievement: evidence from Bangladesh, Appl. Econ. Lett. 12 (2005) 217–221, https://doi.org/10.1080/1350485042000323608.
[44] T. Kohler, Class size and learner outcomes in South African schools: the role of school socioeconomic status, Dev. South Afr. 39 (2022) 126–150, https://doi.org/10.1080/0376835X.2020.1845614.
[45] T. Coupé, A. Olefir, J.D. Alonso, Class size, school size and the size of the school network, Educ. Econ. 24 (2016) 329–351, https://doi.org/10.1080/09645292.2015.1015405.
[46] T. Shen, S. Konstantopoulos, Class size effects on reading achievement in Europe: evidence from PIRLS, Stud. Educ. Eval. 53 (2017) 98–114, https://doi.org/10.1016/j.stueduc.2017.04.001.
[47] E.A. Hanushek, Evidence, Politics, And The Class Size Debate. The Class Size Debate, 2001, pp. 37–65. Available from: http://hanushek.stanford.edu/publications/evidence-politics-and-class-size-debate.
[48] N. Altinok, G. Kingdon, New evidence on class size effects: a pupil fixed effects approach, Oxf. Bull. Econ. Stat. 74 (2012) 203–234, https://doi.org/10.1111/j.1468-0084.2011.00648.x.
[49] Y. Wang, H. Cai, C. Li, Z. Jiang, L. Wang, et al., Optimal caliper width for propensity score matching of three treatment groups: a Monte Carlo study, PLoS One 8 (12) (2013), e81045, https://doi.org/10.1371/journal.pone.0081045.