Evaluating the impacts of the Brazilian model for military public schools.

Jevuks Matheus de Araújo[1]*
Pedro Jorge Holanda Alves[2]
Ana Karolina Acris Melo[3]
Margaria Noélia de Aguiar[4]
Rômulo Eufrosino[5]

Abstract

This article evaluates the impact of the militarization of public schools in the state of Goiás, Brazil, on student performance in standardized tests and age-grade distortion. Using data from the School Census and Prova Brasil between 2007 and 2020 and adopting a difference-in-differences approach with multiple periods (event study), the results indicate that militarization led to a reduction in age-grade distortion and an increase in math, Portuguese, and the Basic Education Development Index (IDEB) scores. The research also examines the impact of militarization on variables related to school violence, revealing a significant decrease in threats to professionals, robberies, and the presence of alcohol, drugs, and weapons among students. This analysis of transmission channels offers insights into how militarization contributes to improving the school environment and academic performance. Robustness analyses, including propensity score matching and the "Honest DiD" method, confirm the significance of the results, even in scenarios where the parallel trends assumption is violated. The results contribute to the debate on educational policies and the expansion of the civic-military school model in the country, providing a valuable benchmark for similar programs across Brazil. However, further research is needed to assess long-term impacts and other relevant dimensions, such as the development of socio-emotional skills and civic education.

Keywords: Educational Policy; Public education; Military education; Educational performance; Impact assessment

JEL classification: I20; I28; I21

[1] Research Fellow at the Laboratory of Economics and Public Policy Evaluation (LEAP-UFPB). Email: jevuks@gmail.com Orcid: https://orcid.org/0000-0002-5618-4502 ***Corresponding author**
[2] Research Fellow at the Laboratory of Economics and Public Policy Evaluation (LEAP-UFPB). Email: pedrojorge963@gmail.com Orcid: https://orcid.org/0000-0001-9340-030X
[3] Research Fellow at the Laboratory of Economics and Public Policy Evaluation (LEAP-UFPB). Email: ana.acris@live.com Orcid: https://orcid.org/0000-0002-9688-1249
[4] PhD student in economics at University of Nebraska–Lincoln (UNL). Email: mnoelia.aguiar@outlook.com Orcid: https://orcid.org/0000-0003-2928-7589
[5] Research Fellow at the Laboratory of Economics and Public Policy Evaluation (LEAP-UFPB). Email: romuloeco@outlook.com Orcid: https://orcid.org/0000-0002-6706-627X

1 Introduction

Military schools are common in many countries. In the USA, the Junior Reserve Officers Training Corps (JROTC) was created in 1916; in India, the Rashtriya Military School was established in 1946. In Brazil, the militarization of public education has become prominent in recent years, especially in 2019 when, in partnership with the Ministry of Education (MEC) and the Ministry of Defense (MD), the National Program of Civic-Military Schools (PECIM)[6] was launched. According to the authorities involved, it is based on public schools that follow the existing militarized method in the country and aims to improve the national indicators of basic education.

Among the 27 federative units that make up the administrative structure of Brazil, the state of Goiás stands out for pioneering the implementation and extension of such a model. The program implemented in the state serves 66,000 students in 60 schools. The objective of this research is to evaluate the impact of the program on student performance in standardized tests (language and mathematics), on the Basic Education Development Index (IDEB), and the age-grade distortion rate.

This process of militarization of the public school system has raised several discussions in the academic field regarding possible impacts on education. On the one hand, some researchers argue that school structures based on the characteristics of a militarized environment, such as discipline, respect for hierarchy, and compliance with a code of conduct, could help in the development of non-cognitive skills, i.e., self-discipline, time management, individual and collective responsibility, the definition of objectives, and teamwork (Price, 2008). Studies show that skills such as these are as important as cognitive skills and could influence educational choices and, consequently, student performance (Heckman, Stixrud, and Urzua, 2006). On the other hand, some authors disagree (Benevides and Soares, 2020; Galaviz et al., 2011), pointing out that this would not contribute to the formation of critical thinking among students.

In the United States, during the 1990s, a set of strict measures to teach compliance with the rules (and indirectly reduce the rates of violence) was adopted in several schools,

[6] However, the program was terminated in 2023. Despite this, the debate over the militarization of public schools continues, particularly after the state of São Paulo launched, in 2024, its own civic-military school program, with implementation starting in 2025, following Goiás' example.

the so-called Zero Tolerance policy. According to the opinion of education professionals, students who fail to comply with pre-established determinations receive very severe punishments (Arum and Ford, 2012). A study conducted by the American Psychological Association Zero Tolerance Task Force (2008) found that excessive school punishments, such as suspending or expelling students, contrary to common sense, did not improve their satisfaction with the institution/school environment. Therefore, the measures adopted by this policy in the United States could not achieve efficient school discipline.

Pema and Mehay (2009) evaluated the JROTC, and their results indicate that students who participated in the program had higher rates of military enlistment; however, they presented lower rates when enrolling in higher education institutions. Before this study, Elliott, Hanser, and Gilroy (2000) analyzed the JROTC Career Academy and, unlike the previous results, the program presented a positive effect on academic performance, especially in dropout rates and school attendance.

There is a scarcity of studies analyzing a militarized pedagogical line, and there is divergence regarding its impacts on the academic performance of students who follow it. Given the number of students assisted by the policy of civic-military schools in Goiás, the results of this article are relevant in the discussion of adopting a system based on the military characteristics of public education. This is because of the absence of studies that empirically analyze the effects of militarization on important educational variables in developing countries.

The results of this research indicate that the process of militarization employed in public schools in Goiás influenced student performance. Considering the canonical difference-in-differences (DD) model, the restructuring of school management resulted in a 10% reduction in age-grade distortion and a significant improvement in the results of standardized tests: an increase of 0.60 in the IDEB score and an increase of 15.25 and 11.61 in the mathematics and Portuguese scores, respectively.

Robustness tests confirm the significance of the results. Using the reweighted event study by Callaway and Sant'Anna (2021), the results also show that the impact is persistent and grows with exposure to treatment. Four years after treatment, it is possible to observe that the effect is much higher than the average one initially observed. Additionally, the results reveal that militarized schools experienced a significant reduction in various variables related to school violence, such as threats to professionals, robberies, and the presence of alcohol, drugs, and weapons among students. Robustness analyses, including the "Honest DiD" method, confirm the significance of the results, even in scenarios where the parallel trends assumption is violated.

The rest of the article is organized as follows. We describe the process of militarization of schools in the next section. In Section 3, we explain the details of the methodology adopted. In Section 4, we present the main results, robustness tests, and heterogeneity exercises. Finally, in Section 5, we present the conclusion and discuss some implications of the policy.

2 Process of militarization of schools in Goiás

The beginning of the militarization of schools in the State of Goiás occurred in 1976, by State Law No. 8,125, which allowed the creation of the first CEPMG (Project State School of The Military Police of Goiás). However, activities were implemented only in 2000. Despite the initial objective of the CEPMG in meeting the needs of families from the State Military Corps[7], the educational actions reached other students. In the initial years, this implementation was slow, until 2012 there were only six schools of this kind in the state. As of 2013, the number of public schools under this system expanded significantly. In 2021, according to the Secretary of Education, Culture and Sports (Seduce), 60 CEPMG are operational in 47 municipalities of the state and serve 66,000 students, from elementary to high school levels.

This process consists of the transfer of administrative responsibilities of the institution to the Military Police, the Fire Department, or state security authorities, while teaching-pedagogical activities continue under the responsibility of education professionals. According to the Military Police, among the objectives of the CEPMG are providing students with a quality environment for the development of school

[7] In Brazil, public security is essentially a responsibility of the states. Each state has two police forces—The Civil Police and the Military Police. The latter plays a different role from the typical Military Police of other countries. In Brazil, they perform tasks in the field of prevention that often overlap with those of other government agencies. They are a uniformed force whose task is to conduct ostensive patrolling and the maintenance of order. Therefore, for the purposes of this study, we will use the terms Military Police and State Police interchangeably. When referring to the militarization of schools, we will be talking about the actions conducted by the State Police in public schools.

activities, as well as contributing to their formation as citizens (becoming aware of their rights and duties) and, consequently, of society itself.

To this end, continuous actions are conducted to qualify professionals (especially teachers), with a focus on improving the teaching-learning process; stimulate the participation of those responsible for the students and of decision-makers, with the entire community involved in the school management. Regarding the process of selecting students, according to what is available in the CEPMG Rules of Procedure, those interested in enrolling must do it in only one of the participating schools since the distribution of vacancies is conducted via lottery.

According to the Seduce and the State Police, the process of militarization of state schools in Goiás aims to identify possible problems involving school violence and ensure the teaching-learning process, boosting the development of basic education in the state through the construction and strengthening of the relationship between educational agents and the community, besides trying to reduce (or even eliminating) the violence in the school environment. However, there are no official sources (documents) that establish the criteria for choosing schools to join the program.

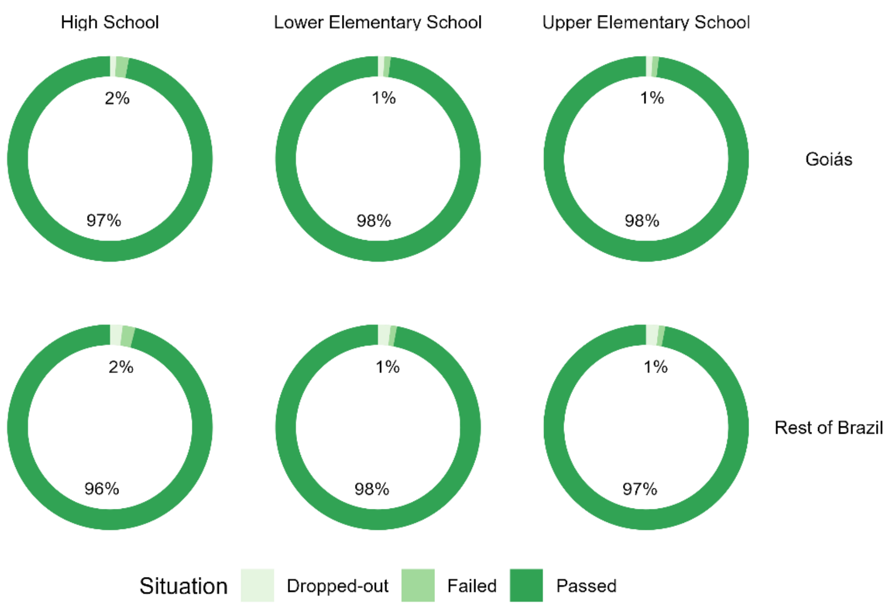
Despite the lack of detailed official information on the selection and adhesion criteria for schools to the CEPMG model in Goiás, a pattern can be observed that suggests the influence of certain factors in the choice of institutions. Schools with a history of low academic performance, high rates of violence or school evasion, located in vulnerable areas and with adequate infrastructure to house the program seem to be prioritized. The community's demand for an education based on discipline and traditional values also emerges as a possible relevant criterion. However, the lack of government transparency regarding the selection process and the absence of official documents detailing the criteria prevent a more precise and in-depth analysis, raising questions about the impartiality and equity in the choice of schools.

2.1 Educational Panorama

The objective of this subsection is to present the educational panorama of schools in the state of Goiás compared with the results of the other states of Brazil and evaluate the temporal trajectory of the schools themselves. An important factor to be evaluated is age-grade distortion[8], which represents the percentage of students enrolled at the wrong age in schools. The main causes for this are evasion (when students give up the current school year but return in the following one), dropping out (when students leave school and do not return in the following year), and failing (when students do not reach the cut-off score and must repeat a school level).

According to Bissoli and Rodrigues (2010), the characteristics of age-grade distortion may be one of the main determinants of school inefficiency. If students are held back, the government needs to generate investment for them over a longer period. According to Figure 1, in the State of Goiás, the percentage of students who dropped out or failed in 2019 was 4% for the initial levels of elementary school, 3% for the final levels of elementary school, and 6% for high school. These results are lower than the rest of the country, which presents 10%, 6%, and 11%, respectively.

Figure 1: Percentage of students who passed, failed, and dropped out by level of education - Brazil and Goiás (2019)

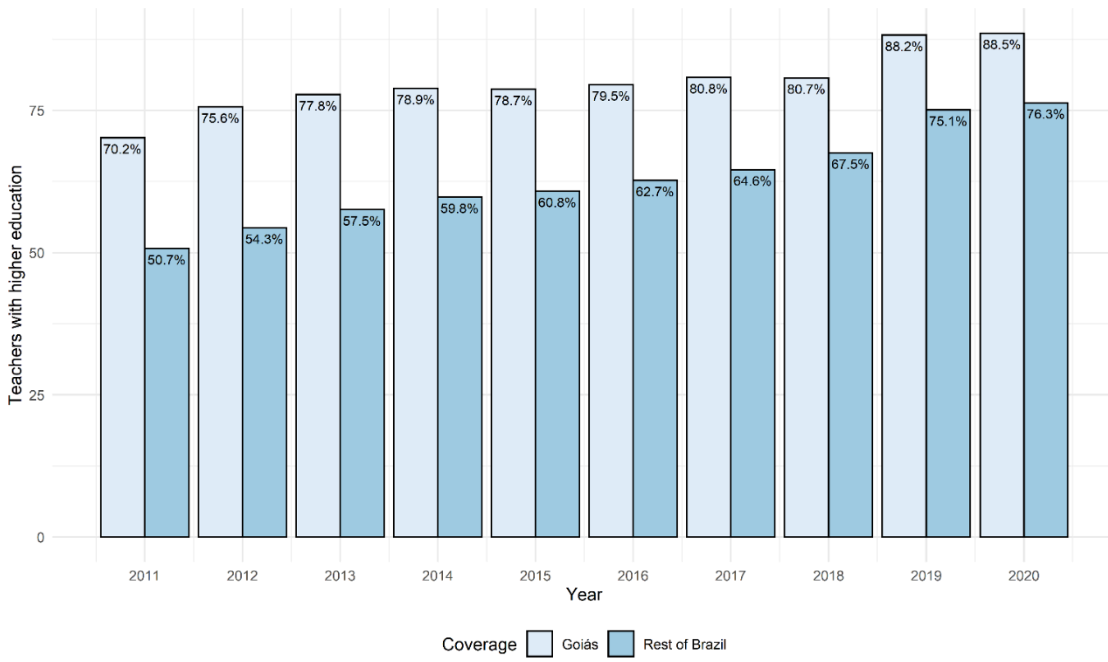


[8] According to the Ministry of Education (MEC), the objective of measuring the age-grade distortion rate is to show the dispersion between the age and the grade that the student is enrolled in. Based on this indicator, one might know the percentage of students which do not fit into the correct age. In Brazil, six years is the appropriate age to enter elementary school, where students will go through a 9-year process (initial and final levels). Moreira et al. (2013) believe that the high age-grade distortion occurs due to the high rate in school evasion, which in turn generates negative impacts on educational performance.

Note: Data collected from the National Institute of Educational Studies and Research Anísio Teixeira (INEP).

What appears in Figure 2 follows the literature explored: schools in the State of Goiás, compared to the rest of Brazil, have higher proportions of teachers with complete qualifications. Data from the School Census indicate that in 2011, in this state, there were an average of 70.2% of teachers with higher education and only 50.7% in Brazil. Between 2011 and 2020, this proportion grew by 26.06% in Goiás, amounting to 88.5% of teachers with higher education in 2020.

Figure 2: Percentage of teachers with higher education - Brazil and Goiás (2011 - 2020)



Note: Data collected from the National Institute of Educational Studies and Research Anísio Teixeira (INEP).

3 Methodology

3.1 Data description[9]

To assess the results presented more robustly, this article seeks to use near-experimental methods to evaluate the impact of the adoption of the program on school performance and dropout rates. Data were collected at the National Institute

[9] Appendix B provides a comprehensive description of the data.

of Educational Studies and Research Anísio Teixeira (INEP) regarding the School Census and Prova Brasil. The first brings detailed information about all schools in Brazil, including the characteristics of their students and teachers. The second is a test applied to public schools aimed at evaluating the quality of the education offered by the Brazilian educational system.

This data set allowed us to build a sample with an unbalanced panel from schools between 2007 and 2020, containing information from before and after the implementation of the program. All control variables with absent data had information imputed with the mean between the schools. We have adopted the year 2013 as the initial one, that is, we have 6 years of pre-treatment and 8 years after it. In Table 1, the descriptive statistics values of our database (Chart A1 – with description attached) are shown, comparing the differences in means between schools that adopted or not the program in the final levels of elementary school. Columns (2) (3) show the Mean and Standard Deviation for the control group, and columns (4) (5) show the same statistics for schools that adopted the program. Column (6) presents the difference in means between columns (2) and (4).

It is noticed that, on average, the results of the performance indicators of students in the last levels are the same as before 2013 (first year of treatment) and, after it, schools that adopted the program became better at some point. When evaluating the results for mathematics and Portuguese, the difference between the averages shows better performance for schools in the program.

Regarding the age-grade distortion rate, the data presented in the descriptive statistics show that the mean difference between groups is significant, indicating that students in this program present a better situation. Considering the school structure, the data indicate that the institutions under the program are different. This evidence provides a preview of the findings, showing that institutions participating in the program perform better.

Table 1: Descriptive Statistics of Control and Treated Schools, Before and After 2013.

	Control		Treated		Mean
	Mean	S.D.	Mean	S.D.	Difference
Panel A. Before 2013					
Number of Students	97.75	167.19	336.70	206.77	-238.95*
% Male students	0.19	0.16	0.22	0.16	-0.03*
% White students	0.52	0.08	0.50	0.03	0.02*
% Students who failed	9.09	9.27	10.60	6.73	-1.51*
% Students who dropped-out	6.34	7.77	8.05	6.83	-1.71*
% Students who passed	84.57	12.25	81.35	10.81	3.22*
The average number of students per class	25.24	8.80	30.64	5.65	-5.40*
Computers for students	7.55	22.44	18.20	10.45	-10.65*
Computers	5.36	14.91	14.16	9.74	-8.80*
Distortion Age series	32.39	14.85	33.27	11.91	-0.88
IDEB	3.71	0.67	3.62	0.68	0.09
Mathematics	238.75	14.61	240.75	11.85	-1.99
Portuguese	232.59	14.99	234.55	13.05	-1.96
Classrooms	6.86	8.59	13.99	7.97	-7.14*
Public transportation	0.20	0.30	0.10	0.10	0.11*
Total Transfers PDDE	7511.30	9200.44	17567.85	12804.74	-10056.55*
Teachers with higher education	83.27	27.62	93.39	7.90	-10.12*
Panel B. After 2013					
Number of Students	80.56	150.31	396.92	242.61	-316.36*
% Male students	0.20	0.15	0.18	0.13	0.02*
% White students	0.52	0.07	0.49	0.04	0.04*
% Students who failed	6.28	8.27	5.40	4.74	0.88*
% Students who dropped-out	2.92	5.14	2.44	3.11	0.48
% Students who passed	90.80	10.18	92.16	6.59	-1.36*
The average number of students per class	24.54	8.80	32.40	5.21	-7.85*
Computers for students	12.69	691.83	13.24	9.13	-0.56
Computers	8.89	29.21	15.87	11.65	-6.98*
Distortion Age series	24.22	13.11	16.54	12.16	7.68*
IDEB	4.84	0.70	5.26	0.85	-0.42*
Mathematics	256.83	16.71	269.06	20.21	-12.22*
Portuguese	256.12	16.32	267.81	19.20	-11.69*
Classrooms	5.31	5.78	9.70	6.70	-4.39*
Public transportation	0.19	0.29	0.08	0.11	0.11*
Total Transfers PDDE	9402.29	69708.92	16956.94	10644.94	-7554.65*
Teachers with higher education	85.68	23.16	89.08	10.75	-3.40*

Note: Data collected from the National Institute of Educational Studies and Research Anísio Teixeira (INEP). The table represents the average and standard deviation of treated and control schools. * Statistically significant

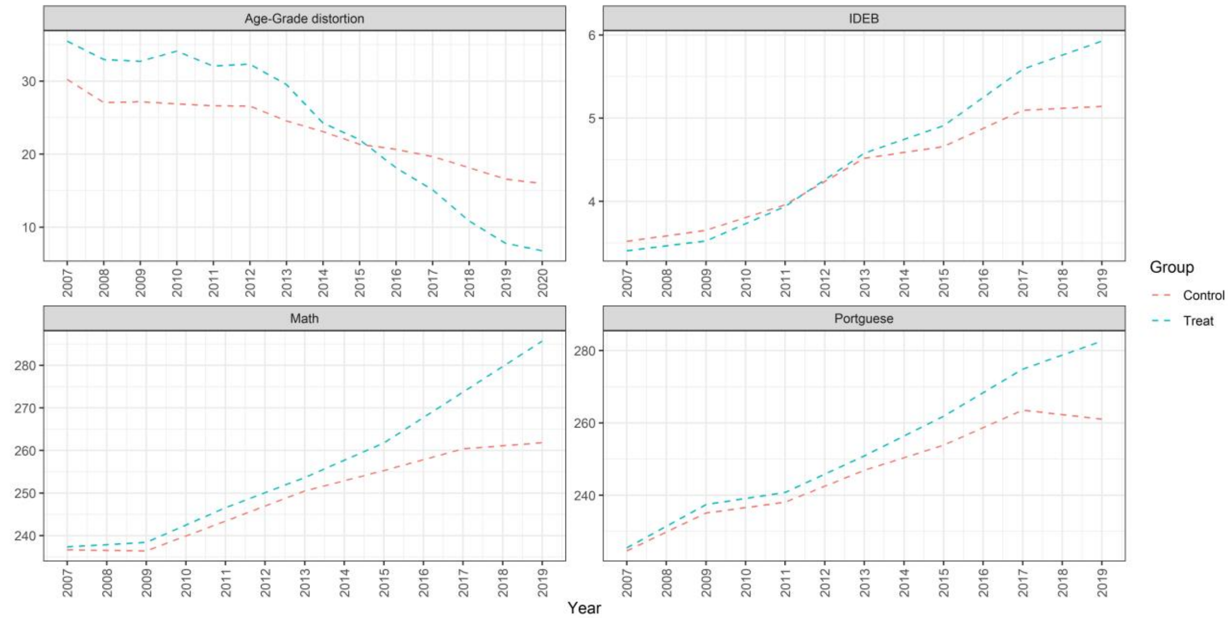
3.2 CHARACTERIZATION OF THE TREATED AND CONTROL GROUPS

To better characterize the groups studied, the treated and control groups will be differentiated according to the parameters studied. Figure 3 illustrates the average trajectory of the main variables of student performance used in this work, such as age-grade distortion, IDEB grade, and the grade in the Brazilian Mathematics and Portuguese Test during the period from 2007 to 2020. Initially, analyzing the evolution of the results of the age-grade distortion in the schools that adopted the military management model (blue) and in the schools that did not adopt it (red), it is possible to observe that in the years that precede the treatment (before 2013), the distortion of students was higher in treated schools. After treatment (after 2013), both groups showed a drop in the age-grade distortion, however, the schools that adhered to the military model showed a more expressive reduction, reaching a number below that of the control group from 2015.

The evolution of the IDEB score during the period of the analysis is shown in Figure 3. It is noted that both groups, control and treated, had similar trajectories and levels during the pre-intervention period. After adopting the policy, in 2013, both continued to show a growing trajectory. However, the schools under the program showed superior performance concerning the other public schools. In addition, it is also possible to observe that from 2017 onwards, student performance from conventional public schools, in the IDEB, began to have a constant trajectory, while those under the military model continued to follow an upward trajectory.

It is also possible to see in Figure 3 that the evolution of the grades in mathematics presents a similar result to those in IDEB. During the pre-treatment period, the treated and control groups followed nearly identical trajectories. As of 2013, student performance from militarized schools became superior to that from regular schools. Finally, regarding the grades in Portuguese and mathematics grades, both groups followed similar trajectories during the pre-treatment period. As of 2013, the performance of the treated group began to differ. After 2017, while students from control schools showed worse performance in Portuguese, the treated group maintained a growth trajectory.

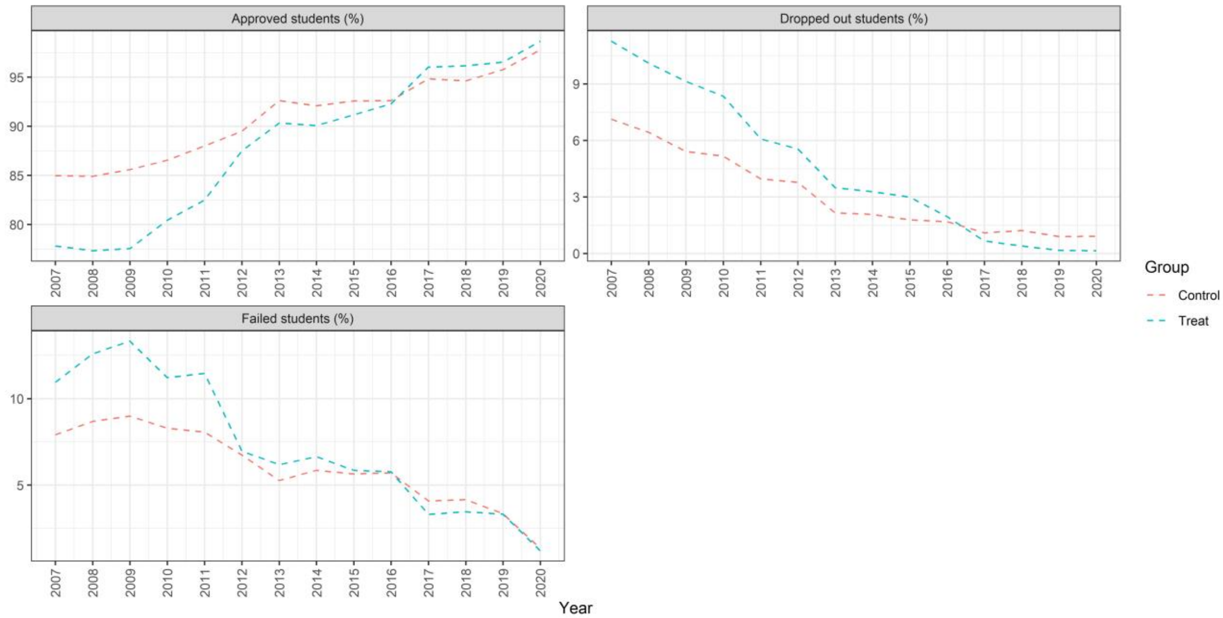
Figure 3. Evolution of the outcome variables for treated and control schools, from 2007 to 2020.



Note: Data collected from the National Institute of Educational Studies and Research Anísio Teixeira (INEP). The graph represents the average of treated and control schools.

A possible problem in the analysis is the bias in terms of the profile of students from schools belonging to the program. Although the new openings are filled by lottery, the schools in the program may be attracting better students and losing those with the worst performances. As we do not have individual student data, we seek to identify characteristics of student profiles in data gathered by schools. Thus, in Figure 4, we can observe school flow data represented by the pass, fail, and drop-out rates. The data show that the behavior of the variables is very similar both in the pre- and post-treatment periods. They show that, despite a reduction in the gap between the two groups, it is not possible to observe that these indicators have any difference between the groups in the post-treatment period. The only exception is the school dropout variable which, for the treated group, had a small inflection between 2015 and 2017. The behavior of the indicators shows that admission by lottery reduces the possible bias of self-selection.

Figure 4: Evolution of outcome variables by groups of treated and control schools of the civic-military program – Goiás, from 2007 to 2020.



Note: Data collected from the National Institute of Educational Studies and Research Anísio Teixeira (INEP). The graph represents the average of treated and control schools.

3.3 Empirical Strategy

The analysis of this work is built around an important regulatory change that occurred in the 2000s in the state of Goiás, in which the government began a process of school militarization, distancing from the civil management of the Departments of Education. Regarding the selection criteria, the chosen institutions are characterized by high rates of violence both in their vicinities and facilities. Thus, this process aims to end violence in the student environment, whether against staff or the students themselves, as well as to improve school performance.

To date, the civic-military schools of Goiás or the Project State School of The Military Police of Goiás (CEPMG), have not been evaluated. Because of the pioneering expansion of the military model in the public network of schools in this state, filling this gap and verifying student performance contributes, in terms of educational policy, to the current debate on the expansion of this model of education throughout Brazil.

Until 2007, the process of implementing military schools occurred more discreetly, with the implementation of six units of civic-military schools. From then on, between 2013 and 2018, there was an accelerated expansion of the process, as can be observed on Table 4, in the Appendices.

Thus, this work will make use of the canonical difference-in-differences (DD) model to estimate the homogeneous effect of the militarization of schools on the outcome variables. As the coverage of the school militarization program did not occur uniformly across the municipalities of Goiás, we took advantage of the staggered way of estimating their causal effect using different strategies that allow estimating the heterogeneous effects of the treatment due to different dates of entry into the program. For this, a reweighted event study proposed by Callaway and Sant'Anna (2021) was used. In addition, sensitivity and robustness analyzes were conducted using pairing techniques and verifying effects on other school flow variables.

3.3 The Difference-in-Differences Estimator (DD): The homogeneous effect of the militarization of schools

Also called a two-way fixed effects estimator, the difference-in-differences (DD) estimator requires information from the treated and control groups, for at least two periods. From this, it will compare groups before and after the intervention. In addition, the method is commonly used when there are heterogeneous unobservable characteristics among groups that influence participation in politics, but that does not vary over time (Meyer, 1995).

This work will initially use the panel data approach with the difference-in-differences estimator, also known as the difference-in-differences estimator with two-way fixed effects (TWFE). As previously stated, this method will allow the comparison of public schools that have undergone the process of militarization with non-militarized ones before and after this process, in the state of Goiás. The model that admits the homogeneous effect of participation (canonical DD model), used to estimate the effectiveness of school militarization on performance indicators, is described in the following equation:

$$y_{it} = \delta_t + \beta Program_{it} + \sum_k \gamma_k X_{it,k} + \psi_i + \epsilon_{it} \quad (1)$$

in which y_{it} represents our outcome variables for school i in year t. $Program_{it}$ is a treatment-indicator variable that equals schools participating in the military civic program after the year it was adopted (if the school participated in the program in 2013, then it will be 1 in 2013 and subsequent years), and, otherwise, zero is assumed. δ_t is a year fixed effects, that adjusts for shocks that are common to all schools at a specific point in time, and ψ_i is a fixed school effect that nonparametrically controls time-invariant establishment factors. Additionally, $X_{it,k}$ is a vector containing our control[10] variables e, and ϵ_{it} is the term of error. β is the parameter of interest, which measures the impact of the program on the outcomes of the schools.

We have also combined the canonical difference-in-differences (DD) method with Propensity Score Matching (PSM). This method proposed by Heckman, Ichimura, and Todd (1997) consists of two steps. First, the PSM is performed and then the DD regression is weighted by the weights assigned to the controls from the propensity scores estimated in the first stage. The PSM enables the pairing of schools that did not participate in the program with those that did it according to similar observable characteristics in the pre-treatment period (base year). The estimated propensity scores, by the PSM, are then used to calculate the weights needed to balance the control group schools under the common support region making them, on average, similar to the treated ones. In the case of this work, the weighting will be done using different matching techniques such as Nearest-Neighbor Matching, Radius Matching, and Optimal Full Matching[2]. With the control group selected by the PSM, one can then estimate the mean effect of the treatment on the treated group using the DD method (Bertrand and Mullainathan, 2004; Ravallion, 2007).

[10] the number of students in school, proportion of male students, proportion of white students, number of computers for students, number of computers and rooms, average number of students per class, percentage of teachers with higher education, number of resources transferred from the FNDE to schools, and a dummy showing whether the school has public transport available for students.

3.4 The Method by Calaway and Sant'Anna (2021): The heterogeneous effect of the militarization of schools

The method of difference-in-differences requires the trajectory of the result variables for the treated and control groups to evolve in parallel before treatment, that is, the hypothesis of parallel trends. According to it, before the period that marks the beginning of the militarization of public schools in Goiás, it is admitted that the treated and control schools have the same change in the result variables over time (Meyer, 1995; Foguel, 2012). In other words, if the mean results for the treated and control groups have followed parallel paths over time in the absence of treatment, the mean effect of treatment for the treated subpopulation (TA) can be estimated by comparing the mean change in the results of the treated group with that of the control group. That is, any effect of the treatment would be captured by the difference in the result variables before and after it (Lechner et al. , 2011).

However, some recent studies show that satisfying the assumption of parallel trends is not enough to establish reliable estimates in environments where treatment time varies and the effects of treatment evolve (Athey and Imbens , 2022 ; Goodman-Bacon , 2021 ; Callaway and Sant'Anna , 2021 ; Sun and Abraham , 2021) this scenario, the Average Treatment Effect (ATE), which is a weighted average of all results among the group-period pairs, may be biased by the presence of negative weights that result from the heterogeneity of the treatment. Therefore, when using the standard difference-in-differences (DD) model with two periods, or the difference-in-difference estimator with two-way fixed effects (TWFE), within an environment with variation in treatment time, the estimator would provide a weighted sum of different treatment effect means, i.e. the TWFE estimator will not capture the Average Treatment Effect on the Treated (ATT) of the policy, but rather a weighted average of all DD estimators, which skews the analysis of the impact of the policy, since heterogeneous treatment effects may be present.

In other words, when there is variation in treatment time, the standard DD coefficient (TWFE) is a weighted average of all standard DD estimators, where each estimate is associated with a difference-in-standard comparison between each treated group with another one, on a specific date. Goodman-Bacon (2021) explains that each of these comparisons is weighted, and some of these weights can be negative because when comparing a unit that will be treated in the future with an already treated one, the latter serves as a control unit, as well as changes in its treatment effect over time are subtracted from the final estimate of the TWFE. In addition, the author states that the TWFE estimator assigns more weight to larger treated groups, that is, the more schools adopt the military education system simultaneously, the more they influence the final aggregate estimate.

However, as Sun and Abraham (2021) point out, the problem of negative weights mentioned earlier can also affect the delays and advances of event-study estimates, even when all treated observations are pooled. Therefore, considering that the public schools in Goiás made the gradual transition of management between 2000 and 2018, and to ensure that the hypothesis of parallel trends is respected, this work also estimates the dynamic effects of the treatment based on the method proposed by Callaway and Sant'Anna (2021), which estimates the average effect of the treatment in the time of the group, assuming that it is possible to satisfy the hypothesis of parallel trends after conditioning on observable variables in the pre-treatment period. In addition, their methodology also allows for estimating the aggregate effects of treatment by relative time (i.e., in the form of an event-study approach), which is the main approach used in this study.

According to Callaway and Sant'Anna (2021), there are several causal parameters of interest when the treatment effect varies between treated groups over time. In this case, the average treatment effects on the treated (ATT) is a function of the treated group g, in which a group is formed according to the period that the units are first treated (e.g., schools that have joined the military management for the first time in 2013 and those who did it in 2015 are in separate groups), and period t. The parameter of interest becomes ATT (g, t), obtained from a two-step estimate with a bootstrap procedure to perform an asymptotically valid inference that adjusts for standard errors for autocorrelation and clustering.

Following the notation by Callaway and Sant'Anna (2021), suppose that there are T periods where t = 1, ..., T and that D_{it} is a binary variable equal to 1 if the unit is treated, and 0 otherwise. G_g is also a dummy variable equal to 1 when a unit is handled for the first time in the period, i.e., it will represent the period that the unit is handled, and C is a dummy variable equal to 1 for units never handled. In the specific case of this work, we will focus on estimating the parameter represented by (5.2), which is a way of aggregating the average treatment effect of the time group to highlight the dynamics of the treatment effect. In that case, the equation is

$$\theta_{es}(e) = \sum_{g \epsilon G} 1\{g + e \leq \tau\} P(G = g | G + e \leq \tau) ATT(g, g + e) \qquad (2)$$

in which $\theta_{es}(e)$ is the average effect of treatment participation for the group of units that were exposed to treatment for exactly "e" periods. Thus, following Callaway and Sant'Anna (2021), event studies are estimated separately for each treatment cohort (defined by the year the school entered military administration) and weighted by the number of investigators in each year of the treatment cohort. This procedure guarantees non-negative weights and can better clarify the dynamic effects of the treatment. It is therefore expected that the Callaway and Sant'Anna (2021) method will generate more accurate estimates, and robust effects of the dynamic treatment effect experienced by schools that went through the militarization management process.

The literature discusses the challenges associated with the parallel trend hypothesis, emphasizing the uncertainty arising from self-selection bias despite strategic and methodological precautions. Rambachan and Roth (2023) argue that estimating the causal parameter of interest becomes difficult and imprecise due to this uncertainty. Moreover, recent studies by Freyaldenhoven et al. (2019), Bilinski and Hatfield (2020), Kahn-Lang and Lang (2020), and Roth (2022) highlight the insufficient power of traditional pretreatment conditional parallel line tests to reject the null hypothesis.

In response to these challenges, we implemented the ‘‘Honest DiD’’ sensitivity analysis proposed by Rambachan and Roth (2023). This procedure involves methods and constraints for robust inference and sensitivity analysis in empirical scenarios where the assumption of parallel trends may not hold. The authors demonstrate that identifying the causal parameter of interest is possible by imposing constraints that ensure post-treatment violations of parallel trends are not "very different" from pre trends. Rather than demanding exact validity of parallel trends, constraints are placed on potential post-treatment differences in trends based on the identified pre trends. The application of the ‘‘Honest DiD’’ allows us to explore and ensure the robustness of our findings.

4 RESULTS AND DISCUSSION

The estimates of the indirect impact of the militarization of public schools in Goiás are displayed in Table 2. In this table, we can see results for different specifications of the canonical DD model, with a homogeneous effect of participation in the program (regardless of the time of exposure to treatment) estimated from Equation (1). Thus, the effectiveness of the militarization of public schools in Goiás is computed by four indicators: i) age-series distortion; ii) grade in mathematics; iii) grade in Portuguese; and iv) grade in IDEB.

As it is shown in Table 2, it is possible to verify the significant effects of the militarization of public schools in Goiás on all the impact variables studied. In addition, the results present the correct direction of the impact of the policy. It is worth noting that some specifications include or not controls for the School Census, which include variables regarding information about students, teachers, and school structure, as well as controls for Prova Brasil, including information about students and parents. Analyzing the results of the models for age-grade distortion, it is suggested that schools adopting military education were responsible for reducing the proportion of students with age-grade distortion by 0.82 standard deviations (SD), according to model (1). After isolating the treatment effect with the addition of covariates, model (2) remained highly significant with a 0.73 SD reduction in the age-series distortion.

Then, analyzing the results for the grades in mathematics, schools with military education increased their grade by 0.92 SD for the model (1), with no data for the School Census and Prova Brasil. In model (2), which tries to isolate only the effect of the policy adopted by the state of Goiás, the military education system increased the school grade to 0.82 SD. Regarding results on grades in Portuguese, model (1) suggests that public schools that adopted the military management system increased their grade by 0.68 SD, without the use of covariates, and increased it by

0.58 SD after adding covariates in model (2).

Still, regarding Table 2, the results of the impact of the militarization process on public schools in Goiás according to the IDEB score suggest that this score increased by 0.79 SD in the model (1), without covariates, and in 0.66 SD after adding covariates in model (2).

The results obtained are similar to those observed in other studies, for example, in Muralidharan and Sundararaman (2011), which evaluated a teacher payment program according to performance in India. They assessed the effects of this program on the performance of schools in mathematics and languages and observed a growth of 0.27 SD in the first and 0.17 SD in the latter. Adnot et al. (2017) evaluated a teacher turnover program that was applied in the District of Columbia in the USA, the results showed an increase of 0.14 SD in reading and 0.21 in mathematics. Barrera-Osorio et al (2022) evaluated a public-private partnership program in Pakistan and observed increases of 0.62 SD in mathematics and 0.59 SD in languages.

Table 2: Estimates of the impact of the Militarization Program for Public Schools in the State of Goiás on different impact indicators (2007-2020). Canonical DD specification (homogeneous average effect)).

	Age-Grade Distortion series		Mathematics		Portuguese		IDEB	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Military civic program (β)	-0.819***	-0.729***	0.927***	0.826***	0.687***	0.584***	0.790***	0.666***
	(0.082)	(0.068)	(0.084)	(0.090)	(0.069)	(0.075)	(0.084)	(0.084)
Observations	20,616	14,414	7,050	4,906	7,050	4,906	7,049	4,906
Number of Schools	1,798	1,666	1,396	1,292	1,396	1,292	1,395	1,292
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: In the table, estimates of difference-in-differences model using the following outcomes are shown: Age-Grade distortion, IDEB, and grades for mathematics and Portuguese. We have used robust standard errors clustered at the schools. The standard deviations are presented in parentheses. Controls include the number of students in school, proportion of male students, proportion of white students, number of computers for students, number of computers and rooms, average number of students per class, percentage of teachers with higher education, number of resources transferred from the FNDE for schools and a dummy showing whether the school has public transport available for students. Successive values of the symbols ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

4.1 Heterogeneity and Robustness Analysis

First, the consistency of the main results of this work, exposed in Table 2, will be tested by combining the Propensity Score Matching (PSM) and the Differences-in-Differences (DD) methods. The adoption of this strategy promises to provide robust estimates of the effect of the program on the adopted impact variables. After exposing the results of this combination, we will test whether the results shown in Table 2 are derived from the change from regular school management to military one, using other impact variables, such as school infrastructure and teacher qualification.

In Table 3, the canonical DD estimates using the control group created from the PSM are revealed. Thus, it is possible to observe that the matching method that presented the best results was the Optimal Full Matching, or total matching. From the control group created by this method, the canonical DD results continue with the expected signs, it is also observed that the militarization of public-school management led to a reduction in age-grade distortion and an increase in both the math and IDEB grades.

Table 3: Results of DD with PSM, 2007-2020

PSM Method	Age-Grade distortion	Mathematics	Portuguese	IDEB
	(1)	(2)	(3)	(4)
KNN	-0.315***	0.285***	0.109	0.193**
	(0.083)	(0.107)	(0.094)	(0.095)
Observations	691	300	300	300
KNN Strong Balanced	-0.319***	0.281**	0.109	0.198**
	(0.080)	(0.107)	(0.093)	(0.094)
Observations	743	316	316	316
Radius	-0.315***	0.285***	0.109	0.193**
	(0.083)	(0.107)	(0.094)	(0.095)
Observations	691	373	373	373
Optimal	-0.425***	0.399***	0.226**	0.268***
	(0.085)	(0.110)	(0.095)	(0.101)
Observations	859	373	373	373

Notes: In the table, estimates of the difference in differences model using the following outcomes are shown: Age-Grade distortion, IDEB, and grades for mathematics and Portuguese. In this specification the control groups were balanced in four methods: KNN, balanced KNN, Radius, and Optimal. We have used robust standard errors clustered at the schools. The standard deviations are presented in parentheses. Controls include the number of students in school, proportion of male students, the proportion of white students, number of computers for students, number of computers and rooms, the average number of students per class, the percentage of teachers with higher education, the number of resources transferred from the FNDE for schools, and a dummy showing whether the school has public transport available for students. Successive values of the symbols ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

It is worth mentioning that all the results presented in Table 3 already include controls in the estimates. In addition, the age-grade, mathematics, and Portuguese distortion coefficients became smaller in relation to those in Table 2, that is, the magnitude of the impact seems to decrease when we use an alternative control group.

Finally, contrary to the main result, in Table 3, public schools that adopted the military management system did not have significant effects on Portuguese scores. Therefore, in general, the results of the program seem not to be sensitive to the choice of the control group.

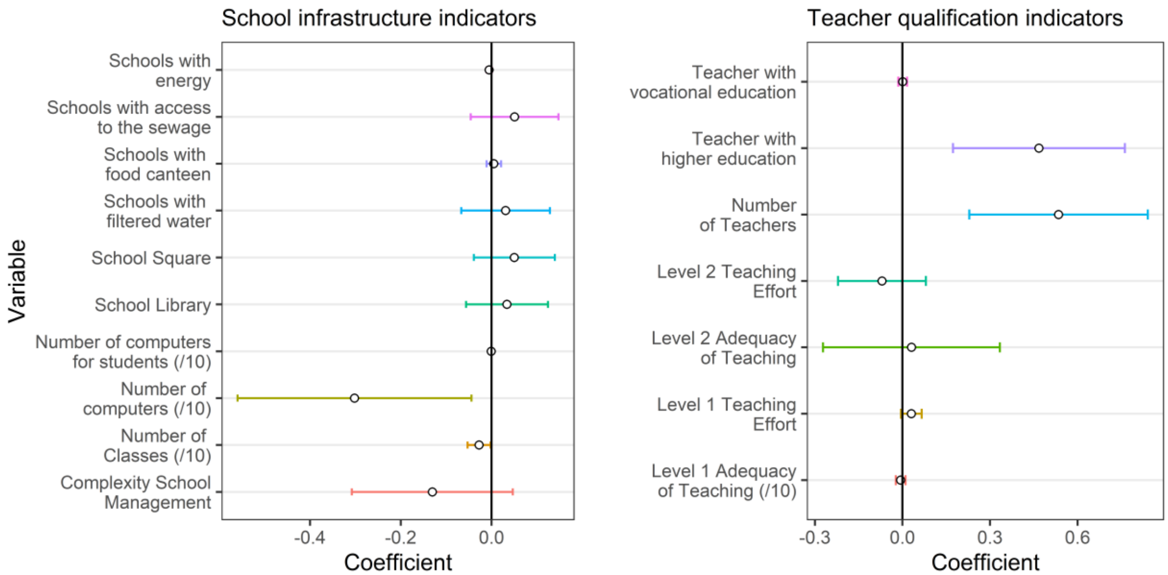
Seeking to deepen the evaluation of the civic-military program in schools, it is first necessary to ensure that the results of the canonical DD shown in Table 2 are derived from the change from school management to the military management model. In other words, the results observed in Table 2 may be related to factors other than the change to the military management model, for example, better performance in the Portuguese and mathematics tests in Prova Brasil, in the IDEB test, and improvements in the age-grade distortion may be related to an increase in the financial resources of the schools that participated in the program. Following the example, if schools have increased their resources, this money may improve their infrastructure or the teacher qualification. In this way, some efforts will be made to guarantee the causality of the results to the civic-military program. Bearing in mind the general objective of the civic-military program in the state, which is to reduce school violence and increase discipline among students, some tests will be conducted to try to fully isolate the effect of the change in military management from other benefits, such as improvement in school infrastructure or teacher qualification.

The impact of military management on school infrastructure variables, also from the canonical difference-in-differences (DD) method, is shown in Figure 5. Regarding the results, it was not possible to observe the effect of military management on these variables. All of them, such as the number of computers per student, number of classrooms, schools with a canteen, or those with a library, showed non-significant coefficients. The same can be seen in Figure 5, with no significant effects of the military management on all teacher qualification variables, such as teaching effort variables (level 1 and level 2) and teacher adequacy (group 1 and group 2)[11]. Thus, as the school infrastructure and teacher qualification did not undergo significant changes during the

[11] For more information on teaching effort variables, see: <https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/docente_esforco/nota_tecnica_indicador_docente_esforco.pdf> and for more information on teacher suitability variables, see: https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/docente_formacao_legal/nota_tecnica_indicador_docente_formacao_legal.pdf.

analyzed period, it is possible to conclude that the results presented in Table 2 are robust, that is, the improvement in student performance was a consequence of the benefits generated by the program.

Figure 5: Effect of the militarization of public schools in the state of Goiás on the teaching and infrastructure indicators



Notes: Estimates of the difference-in-differences model with Multiple Time Periods using the following outcomes are shown: Age-Grade distortion, IDEB, and grades for mathematics and Portuguese. We have used robust standard errors clustered at the schools. The standard deviations are presented in parentheses. Controls include the number of students in school, the proportion of male students, the proportion of white students, the number of computers for students, the number of computers and rooms, the average number of students per class, the percentage of teachers with higher education, number of resources transferred from the FNDE for schools and a dummy showing whether the school has public transport available for students. The standard deviations are presented in 95% of the confidence interval.

Because of the validation of the main results of this work, the goal is now to observe the average effects of treatment at different durations of exposure using the method proposed by Callaway and Sant'Anna (2021). Thus, in Figure 6, the results of the effect of the militarization program on different measures of school performance are shown. It is worth noting that all data were reformulated in relative event time so that we can see the values of the ATT (g, t) parameter expressed there. The estimates include leads and lags for up to four periods before and three periods after the beginning of the militarization process for all impact variables, different only for the age-grade distortion that had leads and lags for 10 periods before and seven after treatment onset[12]. Initially, analyzing the results for the age-

[12] This difference observed in the period of policy analysis between the age-grade distortion and the grades for Portuguese, mathematics and Ideb is justified by the different periodicity of the data. For age-grade distortion, data are observed for the entire period from 2007 to 2020, annually (T0 = 2007,2008,2009,2010,2011,2012; T1

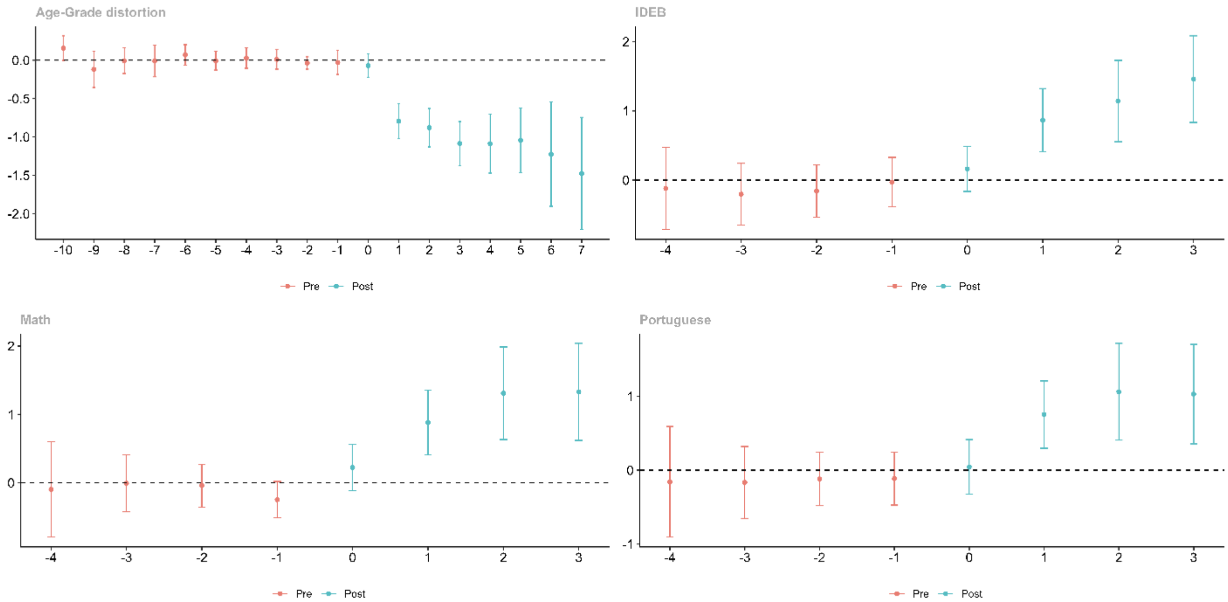
grade distortion in Figure 6, it is possible to affirm that the hypothesis of the parallel lines seems to be confirmed, since all the pre-treatment coefficients are equal to zero, suggesting that there are no differences in the trends of age-grade distortion of control and treated groups before the militarization of public schools began. In addition, the coefficients associated with the years after treatment onset suggest a large reduction in the number of students who are behind. It is worth noting that the militarization policy influenced the age-grade distortion only after the second year of treatment. This effect remains negative and the magnitude of the impact increases over the next seven years.

The estimate for the results of the math and Portuguese tests are also shown in Figure 6. In both cases, it is not possible to observe any anticipated effect of the militarization of public schools in Goiás, and all pre-treatment coefficients are also equal to zero. One year after adopting the military model, the first period after initial exposure to treatment, it was already possible to observe a significant increase in the grades for Portuguese and mathematics in these schools. This treatment participation effect increases with time. Finally, analyzing the impact results of the IDEB test, a positive and significant effect is indicated, similar to previous results, and the effect of this policy seems to be persistent throughout the period of analysis.

Finally, the results of the heterogeneous effect under other flow variables also estimated from the method proposed by Callaway and Sant'Anna (2021) are revealed in Figure 7. In general, the results confirm the absence of effects of the program on the variables that try to capture differences in the profile of students in treated schools. Finally, all results presented in this section corroborate the main one of the homogeneous effects of the militarization process, shown in Table 2.

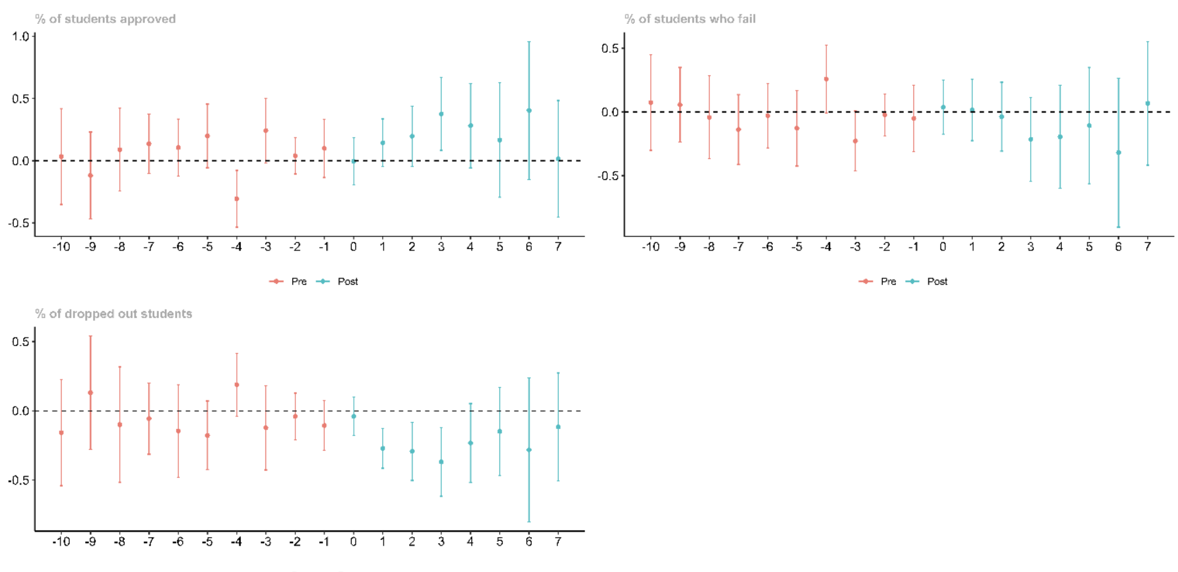
= 2013,2014,2015,2016,2017,2018,2019 ,2020). As for the Portuguese and mathematics grade, obtained from the Prova Brasil, and the IDEB grade, the data are available from 2007 to 2019, however, the frequency is every two years (T0 = 2007,2009,2011; T1 = 2013,2015,2017,2019).

Figure 6: Effect of the militarization of public schools in the state of Goiás on the performance variables studied (2007-2020) - Callaway and Sant'Anna (2021)



Notes: Estimates of the difference-in-differences model with Multiple Time Periods using the following outcomes are shown: Age-Grade distortion, IDEB, and grades for mathematics and Portuguese. The standard deviations are presented in 95% of the confidence interval. No controls are included in these estimations.

Figure 7: Effect of the militarization of public schools in the state of Goiás on the performance variables studied (2007-2020) - Callaway and Sant'Anna (2021).



Notes: Estimates of the difference-in-differences model with Multiple Time Periods using the following outcomes are shown: Approved, Failed, and dropout students. The standard deviations are presented in 95% of the confidence interval. No controls are included in these estimations.

Despite the scarcity of studies in this area, the results for age-grade distortion presented in this article align with existing literature. One component of the distortion indicator is the dropout rate. Authors such as Elliott, Hanser, and Gilroy (2000) and Pema and Mehay (2009), who evaluated similar programs in the United States, demonstrated that students in militarized school models had lower dropout rates. This finding supports the observed reduction in distortion following the adoption of the program in this state.

Regarding the academic performance in the test grade, the results of this study are relatively conflicting when compared to the work by Pema and Mehay (2009). In general, despite positive effects on other variables, such as professional ones, students who participate in the regular Junior Reserve Officers Training Corps (JROTC) have lower academic performance. In this case, these results were expected. One of the justifications used by the authors is that the program would be more aimed at professional life than academic one. That is, any improvement in educational performance would be indirect, due to the development of non-cognitive skills.

On the other hand, the study conducted by Elliott, Hanser, and Gilroy (2000) found that, despite presenting a mild effect, the JROTC Career Academy significantly improved student outcomes. In addition to lower dropout rates and higher school attendance, students in the program achieved higher grades. Similarly, the results of this research align with those of Benevides and Soares (2020), who compared the performance of students from military schools in the state of Ceará, Brazil. When isolating the effect, students in the military education system indeed had higher grades in mathematics, with a learning advantage equivalent to approximately three academic semesters ahead of traditional public school students.

As McPartland (1994) explained, the restructuring and construction of a school environment based on opportunities for success in the academy and on the importance of the school for the community are some of the fundamental tools in encouraging at-risk students to succeed in educational decisions. As in the work by Elliott, Hanser, and Gilroy (2000), a stimulating school environment and the strengthening of the relationship between educational agents and the community may be one of the explanations for the significant improvement in the performance of students from civic-military schools in the state of Goiás.

Pinheiro (2006), from a longitudinal study in high-income environments, observed that children who suffered violence from other students at school may be more likely to miss classes and drop out, which directly affects their educational performance. In the same sense, Devries et al. (2014), in a study conducted in Uganda, associated school violence with poor educational performance, arguing that this violence may be an important factor in poor educational performance in low- and middle-income environments. The results and studies presented so far suggest that teachers and administrators may play an effective role in reducing violent acts in schools and, in turn, improving student performance, which may be the mechanisms that guide the results of this study (Mertoglu, 2015).

4.2 Transmission Channels: School Violence

Another way to ensure the robustness of our main results is to identify mechanisms through which civic-military schools affected students' academic performance. As discussed in Section 2, the central objective of the civic-military schools' program was to reduce or eliminate violence in the school environment. Therefore, we believe that the schools that participated in the program experienced a reduction in problems related to violence within the units, which led to an improvement in students' educational performance.

In this section, we will verify whether civic-military schools influenced some variables that capture the presence of violence within schools. Therefore, our unit of study ceases to be students and becomes schools. To verify whether there has been a reduction in cases of violence in schools, this exercise will use questionnaires applied to principals, made available by National Assessment of Basic Education (Sistema Nacional de Avaliação da Educação Básica—SAEB)[13]. The variables used in this section are a binary survey response from school principals to questions about the occurrence of incidents of violence on and

[13] For more information on SAEB, its objective, methodology, and scale, please see http://ces.ibge.gov.br/en/base-de-dados/metadados/inep/sistema-nacional-de-avaliacao-da-educacao-basica-saeb

around the school campus against students, teachers, or school staff.

Exposure to violence has a strong negative impact on academic outcomes (Beland and Kim 2016; Fry et al. 2018; Kim et al., 2020). This scenario will guide our analysis, taking into account the relationship between the students' environment and their exposure to adverse experiences, such as violence related to robberies, the presence of weapons, drugs, drinks, and their academic performance. Testing this transmission channel is pertinent for several reasons. The first reason is directly related to the criteria for choosing the program itself. The Goiás State Department of Education (Seduce/GO) implemented militarization mainly in schools located in vulnerable regions, with high crime rates and low academic performance. Furthermore, existing literature on the topic of violence in schools states that violent incidents can affect the educational outcomes of the school system directly through the psychological stress imposed on both victims and witnesses (Brown & Taylor, 2008; Liew et al., 2008; Ammermüller, 2012; Sharkey et al., 2012).

From left to right, Table 4 reveals the results for the occurrence of professionals who are victims of violence, the occurrence of threatened professionals[14], the occurrence of robbery, the occurrence of students with alcoholic drinks, the occurrence of students carrying drugs, and the occurrence of students carrying weapons. Except for the variable that captures the occurrence of professionals being victims of violence, the results of the difference-in-differences model show that schools that adhered to military management observed a statistically significant reduction for all school violence variables.

In more detail, a significant reduction of 0.304 in the occurrence of threatened professionals was observed, indicating that military management can have a deterrent effect on threats directed at school professionals. The 0.235 decrease in the incidence of robbery suggests that the presence of more rigid and disciplined management can contribute to a safer school environment. The 0.354 reduction in the frequency of alcoholic beverages among students, followed by the 0.294 reduction in occurrences related to drug possession and the 0.201 reduction in the presence of weapons, reflects the potential effectiveness of control and surveillance policies implemented by military schools. These results corroborate the hypothesis that a more controlled and disciplined school environment can lead to a significant reduction in violent and risky behaviors. However, it is important to note that the variable

[14] These professionals can be any school employee: teachers, principals or school staff.

related to the occurrence of professionals being victims of violence did not show a statistically significant reduction, which may indicate the need for additional specific measures to protect school professionals.

These results reveal the direct mechanism by which the militarization of schools affected students' academic performance. The reductions observed in school violence variables in schools that adopted military management suggest that more rigorous control and security policies may be effective in mitigating these adverse impacts. In conclusion, the implementation of civic-military schools has proven to be a promising strategy for reducing school violence and improving the educational environment, contributing significantly to the improvement in students' academic performance observed in the previous sections.

Table 4: Results of Canonical DD, 2007-2020

	(1) Occurrence of professionals who are victims of violence	(2) Occurrence of threatened professionals	(3) Occurrence of robbery	(4) Occurrence of students with alcoholic drinks	(5) Occurrence of students carrying drugs	(6) Occurrence of students carrying weapons
Tratament	-0.055	-0.304***	-0.235***	-0.354***	-0.294***	-0.201***
	(0.069)	(0.059)	(0.061)	(0.076)	(0.080)	(0.073)
Observations	4,833	4,833	4,833	4,833	4,833	4,833
R-squared	0.552	0.276	0.278	0.201	0.086	0.080
Number of schools	1,494	1,494	1,494	1,494	1,494	1,494

Note: Successive values of the symbols ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

4.3 Sensitivity Analysis: “Honest DiD”

Given the possibility of violations to the parallel trends assumption due to the self-selection of students into civic-military schools, we conduct a sensitivity analysis using the "Honest DiD" approach proposed by Rambachan and Roth (2023). This method allows us to assess the robustness of our findings even when the parallel trends assumption might not hold perfectly.

We focus on four key outcome variables: age-grade distortion and scores in IDEB, Portuguese, and mathematics. For each variable, we estimate the treatment effects for each period, using 95% confidence intervals. The full results are available in the Appendix C (Supplementary Material - Technical Appendix).

Overall, the results demonstrate that the estimated effects are robust to varying degrees of violation of the parallel trends assumption, particularly in the initial periods after the intervention. For instance, in Figure C.1 (age-grade distortion), even when imposing a magnitude constraint of 1 (allowing post-treatment violations to be as large as the largest pre-treatment violation), the treatment effect remains significant in 2014. Similar results are observed for the other outcome variables.

The "Honest DiD" analysis also allows us to assess robustness to different degrees of non-linearity in the violations of parallel trends. In Figure C.2 (age-grade distortion), the treatment effect remains significant even when allowing for a moderate degree of non-linearity (M=0.15). Again, similar results are found for the other variables.

In conclusion, the "Honest DiD" sensitivity analysis strengthens the validity of our results, indicating that the estimated effects of the militarization of schools in Goiás are robust to potential violations of the parallel trends assumption, especially in the initial post-intervention periods. This robustness increases our confidence in the findings regarding the positive impact of militarization on student achievement.

5 Conclusion

This study investigated the effects of the militarization of public schools in the state of Goiás on student performance in standardized tests and age-grade distortion. This research contributes to the national literature as it is the first to verify the impacts that the transition from usual school management to a military one may have on student performance in Brazil, more specifically in this state.

Throughout the development of this work, there was the concern of reducing as much as possible the bias of the calculated estimates, so the first step was to estimate the impact of the policy from the canonical difference-in-differences (DD) estimator, to obtain the homogeneous effect of the militarization of public schools on the performance indicators of the students. In addition, to give greater robustness to the main result of the work and obtain more reliable estimates, the event study strategy was also used, which allows the capture of the effect of this management for each treatment period.

The main results show that the militarization of public schools in Goiás could reduce the age-grade distortion and increase the grades in standardized tests and in the IDEB. Additionally, it was observed that the program led to a significant reduction in various variables related to school violence, such as threats to professionals, robberies, and the presence of alcohol, drugs, and weapons among students. This suggests that the military model can be an effective tool to improve educational outcomes and reduce school violence in the Brazilian context.

Although this study presents limitations, such as the lack of data at the student level to control for individual selection bias in participating in the school lottery for admission or in deciding to leave school, the results achieved are robust and relevant for the debate on educational policies and the expansion of the civic-military school model in the country. The successful experience observed in Goiás serves as a good parameter for the Civic-Military Schools Programs in Brazil, which have been expanded since 2019 and aim at the improvement and restructuring of school management. However, it is important to emphasize that this program is not an exclusive proposal, but rather one of the options to be considered in the wide scope of educational programs and policies that seek to improve the quality of education in Brazil.

In conclusion, this study provides evidence that the militarization of public schools in Goiás had a positive impact on student performance and a reduction in school violence. However, further research is needed to assess long-term impacts and other dimensions, such as the development of socio-emotional skills and civic education, as well as to deepen the discussion about the mechanisms by which militarization affects the school environment and student performance.

References

Ammermüller, A. (2012). Violence in European schools: A widespread phenomenon that matters for educational production. *Labour Economics*, **19**(6), 908–922.

Arum, R., & Ford, K. (2012). How Other Countries" Do Discipline". *Educational Leadership*, *70*(2), 56-60.

Athey, S., & Imbens, G. W. (2022). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*, *226*(1), 62-79.

Beland, L. P., & Kim, D. (2016). The efect of high school shootings on schools and student performance. Educational Evaluation and Policy Analysis, 38(1), 113–126.

Benevides, A. D. A., & Soares, R. B. (2020). Diferencial de desempenho de alunos das escolas militares: o caso das escolas públicas do Ceará. *Nova Economia*, *30*, 317-343.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. *The Quarterly journal of economics*, *119*(1), 249-275.

Bilinski, A. and Hatfield, L. A. (2020), "Nothing to See Here? Non-inferiority Approaches to Parallel Trends and Other Model Assumptions" arXiv:1805.03273 [stat.ME].

Brown, S., & Taylor, K. (2008). Bullying, education and labour market outcomes: Evidence from the National Child Development Study. *Economics of Education Review*, **27**(4), 387–401.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200-230.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Devries, K. M., Child, J. C., Allen, E., Walakira, E., Parkes, J., & Naker, D. (2014). School violence, mental health, and educational performance in Uganda. *Pediatrics*, *133*(1), e129-e137.

Deole, S. S. (2018). Human capital consequences of violence in schools: Estimating the impact of violence in schools on education outcomes in Brazil. *Review of Development Economics*, **22**(1), 287–310.

Elliott, M. N., Hanser, L. M., & Gilroy, C. L. (2000). *Evidence of positive student outcomes in JROTC career academies*. RAND CORP SANTA MONICA CA.

Fry, D., Fang, X., Elliott, S., Casey, T., Zheng, X., Li, J., et al. (2018). The relationships between violence in childhood and educational outcomes: A global systematic review and meta-analysis. Child Abuse and Neglect, 75, 6–28.

Freyaldenhoven, S., Hansen, C. and Shapiro, J. (2019), “Pre-event Trends in the Panel Event-study Design”, American Economic Review, 109, 3307–3338.

Liew, J., McTigue, E. M., Barrois, L., & Hughes, J. N. (2008). Adaptive and effortful control and academic self-efficacy beliefs on achievement: A longitudinal study of 1st through 3rd graders. *Early Childhood Research Quarterly*, **23**(4), 515–526.

Lima, L., FOGUEL, Peixoto, B., PINTO, C. C. D. X., M. N., BARROS, R. P., & MENEZES FILHO, N. (2012). *Avaliação econômica de projetos sociais*. Fundação Itaú Social.

Galaviz, B., Palafox, J., Meiners, E. R., & Quinn, T. (2011). The militarization and the privatization of public schools. *Berkeley Review of Education*, *2*(1).

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254-277.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605-654.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, *24*(3), 411-482.

Kahn-Lang, A. and Lang, K. (2020), “The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications”, Journal of Business and Economic Statistics, 38, 613–620.

Kim, Y. K., Sanders, J. E., Makubuya, T., & Yu, M. (2020). Risk factors of academic performance: Experiences of school violence, school safety concerns, and depression by gender. In *Child & Youth Care Forum* (Vol. 49, pp. 725-742). Springer US.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics*, *4*(3), 165-224.

McPartland, J. M. (1994). Dropout prevention in theory and practice. In: *Schools and Students at Risk. Context and Framework for Positive Change.*

Mertoglu, M. (2015). The role of school management in the prevention of school violence. *Procedia-Social and Behavioral Sciences*, *182*, 695-702.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics*, *13*(2), 151-161.

Pema, E., & Mehay, S. (2009). The effect of high school JROTC on student achievement, educational attainment, and enlistment. *Southern Economic Journal*, *76*(2), 533-552.

Pinheiro, P. S. (2005). UN Secretary-General’s study on violence against children.

Price, H. B. (2008). *Mobilizing the community to help students succeed*. ASCD.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 5, 2555–2591.

Ravallion, M. (2007). Evaluating anti-poverty programs. *Handbook of development economics*, *4*, 3787-3846.

Roth, J. (2022), “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends”, American Economic Review: Insights, 4, 305–322.

Sharkey, P. T., Tirado-Strayer, N., Papachristos, A. V., & Raver, C. C. (2012). The effect of local violence on children's attention and impulse control. *American Journal of Public Health*, **102**(12), 2287–2293.

Skiba, R., Reynolds, C. R., Graham, S., Sheras, P., Conoley, J. C., Garcia-Vazquez, E., ... & Palomares, R. (2006). Are Zero Tolerance Policies Effective in the Schools? An Evidentiary Review and Recommendations An official report of the APA.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175-199.

APPENDIX A

Table A1: List of the Goiás Military Police State College Project (CEPMG) by year of entry

Municipalities	Schools	Year of entry
Goiânia	CPMG Vasco dos Reis	2000
Goiânia	CPMG Hugo de Carvalho Ramos	2000
Rio Verde	CPMG Carlos Cunha Filho	2002
Itumbiara	CPMG Dionária Rocha	2002
Goiânia	CPMG Ayrton Senna	2002
Anápolis	CPMG Dr. Cézar Toledo	2005
Goiás	CPMG Prof. João Augusto Perillo	2013
Jataí	CPMG Nestório Ribeiro	2013
Quirinópolis	CPMG Dr.Pedro Ludovico	2013
Porangatu	CPMG Tomaz Martins da Cunha	2013
Novo Gama	CPMG José de Alencar	2013
Valparaíso	CPMG Fernando Pessoa	2013
Anápolis	CPMG Gabriel Issa	2013
Inhumas	CPMG Manoel Vilaverde	2013
Goianésia	CPMG José Carrilho	2013
Ap. de Goiânia	CPMG Nader Alves dos Santos	2014
Jussara	CPMG Maria Tereza G. N. Bento	2014
Palmeiras	CPMG Cabo Edmilson de S.Lemes	2014
Goiânia	CPMG Míriam Benchimol Ferreira	2015
Goiânia	CPMG Waldemar Mundim	2015
Goiânia	CPMG Jardim Guanabara	2015
Ap. de Goiânia	CPMG Colina Azul	2015
Ap. de Goiânia	CPMG Mansões Paraíso	2015
Ap. de Goiânia	CPMG Madre Germana	2015
Senador Canedo	CPMG Pedro Xavier Teixeira	2015
Formosa	CPMG Domingos de Oliveira	2015
Itaberaí	CPMG Maria Heleny Perillo	2015
Ceres	CPMG Hélio Veloso	2016
Itauçu	CPMG Itauçu	2016
Goiatuba	CPMG Goiatuba	2016
Catalão	CEPIPMG Dr. Tharsis Campos	2016
Posse	CPMG Dom Prudêncio	2016
Caldas Novas	CPMG Nivo das Neves	2016
Goiânia	CPMG Major Oscar Alvelos	2016
Jaraguá	CPMG Sílvio de Castro Ribeiro	2017
Anápolis	CPMG Arlindo Costa	2017
Goianira	CEPMG José Silva Oliveira	2017
Morrinhos	CEPMG Xavier de Almeida	2017
Itapaci	CEPMG Geralda Andrade Martins	2017

Pires do Rio	CEPMG Professor Ivan Ferreira	2017
Iporá	CEPMG Ariston Gomes da Silva	2017
Nerópolis	CEPMG Doutor Negreiros	2017
Goianápolis	CEPMG Benedita Brito de Andrade	2018
Sanclerlândia	CEPMG 5 de janeiro	2018
Ipameri	CEPMG José Pio de Santana	2018
São Luís de M.Belos	CEPMG Américo Antunes	2018
Pirenópolis	CEPMG C. Christóvan de Oliveira	2018
Itapuranga	CEPMG Dep. José Alves de Assis	2018
Luziânia	CEPMG de Luizânia	2018
Rubiataba	CEPMG de Rubiataba	2018
Trindade	CEPMG Castelo Branco	2018
Trindade	CEPMG José dos Reis Mendes	2018
Trindade	CEPMG Pedro Ludovico Teixeira	2018
Uruaçu	CEPMG Francisco Antônio de Azevedo	2018
Guapó	CEPMG Dr. José Feliciano Ferreia	2018
Anicuns	CEPMG Rosa Turisco de Araújo	2018
Alexânia	CEPMG 13 de maio	2018
Hidrolândia	CEPMG Augusta Machado	2018
Goianira	CEPMG Padre Pelágio	2018
Bom Jesus de Goiás	CEPMG Pastor José Antero Ribeiro	2018

Note: The list of schools was made available by the Secretary of State for Education of the Government of the State of Goiás

Appendix B – Description of the Data

This appendix provides a comprehensive description of the data used in the analysis of civic-military schools (ECMs) in Brazil, incorporating various sources of information to ensure a detailed understanding.

B.1 Administrative Data - General Information:

The primary data sources are administrative records from the Ministry of Education (MEC) and the Ministry of Defense (MD), which include detailed information about the establishment, operation, and performance of civic-military schools. These records are mandatory for all schools participating in the civic-military program.

We merge data from military schools with data provided by the School Census and Prova Brasil, which has primary information from schools and is collected by the National Institute of Educational Studies and Research Anísio Teixeira (Inep). The Censo Escolar is a crucial data collection effort that offers extensive insights into Brazil's educational system. Managed by INEP, this census collects data from all public and private schools in Brazil, covering various aspects such as student demographics, school infrastructure, teacher qualifications, and school management practices.

Prova Brasil is a standardized assessment conducted in Brazilian public schools to evaluate students' educational attainment in key subjects. This exam is an integral part of the broader national assessment system and plays a crucial role in the Basic Education Development Index (IDEB), which measures the quality of education in the country. The assessment aims to gauge the proficiency of students in Portuguese and Mathematics, administered to 5th and 9th graders.

B.2. Variable Construction:

The Brazilian School Census, known as Censo Escolar, is the most extensive survey of the Brazilian educational system. Conducted annually by the National Institute for Educational Studies and Research Anísio Teixeira (INEP), it collects data from all public and private schools in the country, offering a comprehensive census and annual overview of the state of education in Brazil.

The process of data collection for the School Census is highly systematic and involves multiple stages. Initially, INEP provides training to local education authorities on the data collection procedures and the use of the Educacenso system. This training ensures that data collection is standardized across all regions, enhancing the reliability and comparability of the data. The data is collected through Educacenso, an online platform where schools input their information directly, ensuring real-time data submission and immediate access to the database.

The schools are required to submit their data through Educacenso during a designated period, typically between May and August. The data includes detailed information about students, teachers, school infrastructure, and management practices. After the submission period ends, the data undergoes a rigorous verification process. INEP uses automated systems to flag inconsistencies and anomalies in the data. Schools are then required to review and correct these issues, ensuring the accuracy and completeness of the data before it is finalized.

E Compliance with the School Census is mandatory for all educational institutions in Brazil. The Ministry of Education enforces this requirement through various means, including linking census participation to the receipt of federal funds and resources. Schools that fail to comply with the census requirements risk losing critical funding and support, which ensures a high level of participation and data accuracy. The School Census is separated into four data tables: school, class, students and teachers.

The student questionnaire gathers demographic details, such as age, gender, race, and disability status, along with academic performance indicators like grades and attendance. The school questionnaire focuses on infrastructure and administrative practices, including the number of classrooms and the presence of school councils. Teacher questionnaires collect data on qualifications, experience, and teaching practices, while the principal questionnaire examines leadership and management activities. While much of this information is collected comprehensively, certain sensitive data remains confidential, with public datasets restricting access to specific details to protect privacy.

B.2.1 Prova Brasil:

Prova Brasil is a standardized assessment conducted in Brazilian public schools to evaluate students' educational attainment in key subjects. This exam is an integral part of the broader national assessment system and plays a crucial role in the Basic Education Development Index (IDEB), which measures the quality of education in the country.

Designed to assess proficiency in Portuguese and Mathematics, the exam is administered to students in the 5th and 9th grades of elementary school. It evaluates their knowledge and skills according to national curriculum standards. Prova Brasil is conducted biennially, typically in October and November, allowing for the capture of students' educational progress over the school year.

The objective of Prova Brasil is to employ a census-based approach rather than a sampling method, meaning it aims to reach all students in the specified grades within public schools across the country. However, logistical and practical constraints sometimes necessitate focusing at least on a representative sample of schools. The selected schools are chosen based on their inclusion in the national educational registry, ensuring wide and representative coverage.

The exam consists of multiple-choice questions designed to measure students' understanding and application of key concepts in Portuguese and Mathematics. The questions align with the national curriculum and are developed to assess a range of cognitive skills, including knowledge recall, comprehension, application, and problem-solving abilities. The administration of Prova Brasil is standardized across all participating schools to ensure consistency and reliability. Trained proctors oversee the exam process, ensuring that all students take the exam under similar conditions in a controlled environment to minimize any external influences that could affect student performance.

Upon completion, the answer sheets are collected and sent to a central processing center where the data is digitized and analyzed using advanced statistical methods. The results generate detailed reports on student performance at school, municipal, state, and national levels. Based on this information, INEP separates the results into questionnaires related to school, class, teacher, student, and school director.

The student questionnaire for the Prova Brasil is designed to collect extensive demographic and academic information. This includes basic data such as age, gender, race, and disability status, as well as detailed information on students' academic progress, including their grades on standardized tests and participation in educational programs. Additionally, the questionnaire gathers information about the students' families, such as household income and parental education levels, providing a comprehensive understanding of the socio-economic context in which students reside.

Focusing on the physical and administrative aspects of educational institutions, the school questionnaire includes questions about infrastructure, such as the number of classrooms, availability of laboratories and libraries, and the condition of sports facilities. It also addresses administrative practices, including the existence of school councils, the implementation of educational projects, and the overall governance structure of the school.

Detailed information about the teaching workforce is collected through the teacher questionnaire, which includes data on teachers' qualifications, years of experience, and participation in professional development activities. It also gathers information on teaching practices, such as the subjects taught and the instructional methods used. This data is critical for identifying the strengths and weaknesses of the teaching workforce and pinpointing areas where additional training and support may be necessary.

The principal questionnaire emphasizes the leadership and management practices of school principals. It includes questions about their educational background, professional experience, and involvement in school management activities. Additionally, it explores the challenges faced by principals and their strategies for overcoming these challenges, providing valuable insights into the effectiveness of school leadership and aiding in the development of programs to support and enhance school administration.

B.1.2 Description of variables:

Key variables constructed from the administrative data include:

Student characteristics. We generated some variables that are related to school students. The first is the number of Students, representing the total number of students enrolled at a given school. It is a fundamental metric that provides insight into the scale of the school's operations and student body size. The data is sourced from the Brazilian School Census, ensuring comprehensive coverage and accuracy. From the number of enrollments we also generate the percentage of male Students and percentage of white Students.

Student performance. From data from educational statistics, it was also possible to extract some information about student performance, as the percentage of students who failed, dropped out or approved a year. High failure or dropped rates can signal issues in teaching quality, curriculum effectiveness, or student support systems. On the other hand, high approval rates indicate effective teaching methods, supportive learning environments, and overall student engagement. From these statistics, it is also possible to create the age-grade distortion rate, which measures the discrepancy between students' ages and their appropriate grade levels. High distortion rates indicate that many students are either over-aged or under-aged for their grade, which can affect their learning experience.

In addition, with the Prova Brasil data, we can also extract some information about student performance, as IDEB Scores and student performance on tests. The IDEB (Basic Education Development Index) score is a composite index that evaluates the quality of education in Brazilian schools. It includes data on student performance in standardized tests (Prova Brasil) and student progression rates. IDEB scores are used to monitor and assess educational outcomes, guiding policy decisions and educational reforms. Furthermore, we can also separate and visualize the performance of school students in each of the tests

Characteristics of schools. School census data allows us to extract a large set of information about schools. The average number of students per class, the objective of which is to understand class size and its impact on learning outcomes. Smaller classes are often associated with more personalized attention and better student performance. Data on the availability of computers for students and the total number of computers at school provide information about access to technology in schools, which is crucial for modern education. Finally, we also collected some variables directly linked to schools, such as the number of classrooms and availability of public transport for children.

In the context of Prova Brasil, variables related to school violence are report by school director and are critical for understanding and addressing safety issues within educational institutions. These variables include the frequency and types of violent incidents reported by students and staff, such as physical fights, bullying, theft, and verbal abuse. Additionally, data is collected on the presence of security measures, such as school guards, surveillance cameras, and protocols for handling violent situations.

Characteristics of Teachers. The percentage of teachers with higher education qualifications is an important indicator of teacher quality. Higher educational qualifications are generally associated with better teaching practices and improved student outcomes.

Chart B1 - Description of the variables

Variable	Description	Source
Number of Students	Total students enrolled in the final years of elementary school at school	Brazilian school Census
% Boy students	Percentage of male students in the final years of elementary school at the school	Brazilian school Census
% White students	Percentage of students declared white in final years of elementary school at school	Brazilian school Census
% Fail Students	Percentage of students who failed a year in the final years of elementary school at school	Brazilian school Census
% Drop-out Students	Percentage of students who drop-out in the final years of elementary school at school	Brazilian school Census
% Approved Students	Percentage of students who approved a year in the final years of elementary school at school	Brazilian school Census
Average number of students per class	Average number of students per year class in the final years of elementary school at school	Brazilian school Census
Computers for students	Number of computers available for students to use at school	Brazilian school Census
Computers	Number of computers at school	Brazilian school Census
Distortion Age series	Age-grade distortion rate of students in the final years of elementary school	Brazilian school Census
IDEB	IDEB grade in the final years of elementary school	Prova Brasil test
Math	School performance in the brazil math test in the final years of elementary school	Prova Brasil test
Portuguese	School performance in the brazil Portuguese test in the final years of elementary school	Prova Brasil test
Classrooms	Number of classrooms at school	Brazilian school Census

Public transportation	Indicative if the school has buses available to transport children	Brazilian school Census
Total Transfers PDDE	Quantitative resource transferred to schools via PDDE	National Education Development Fund
Teachers with higher education	Percentage of teachers with higher education	School Census

Appendix C - Technical Appendix: Sensitivity Analysis Using "Honest DiD" with Smoothness and Magnitude Constraints

The "Honest DiD" approach developed by Rambachan and Roth (2023) provides a framework for conducting robust inference in difference-in-differences (DID) and event study designs when the parallel trends assumption might not hold perfectly. Instead of requiring exact parallel trends, the method imposes restrictions on the post-treatment deviations from parallel trends relative to the pre-treatment differences in trends (pre-trends). Under these restrictions, the causal parameter of interest is partially identified.

Formally, let β be the vector of event-study coefficients, which can be decomposed into:

$$\beta = \tau + \delta$$

where τ represents the treatment effects of interest and δ represents the bias due to the difference in trends between the treatment and control groups. The parallel trends assumption posits that $\delta post=0$, where δpost is the difference in trends after the treatment. However, Rambachan and Roth (2023) propose relaxing this assumption, allowing δpost to be nonzero but imposing constraints on how much it can differ from the pre-trends, δpre.

Two main types of constraints are considered:

1. **Relative Magnitude Constraints (RM(M)):** These constraints limit the magnitude of post-treatment violations of parallel trends relative to the magnitude of pre-trends. For instance, the restriction RM(1) imposes that post-treatment violations cannot be larger than the largest pre-treatment violation.
2. **Smoothness Constraints (SD(M)):** These constraints restrict the extent to which the slope of the difference in trends can change between consecutive periods. The restriction SD(0) imposes that the difference in trends is exactly linear.

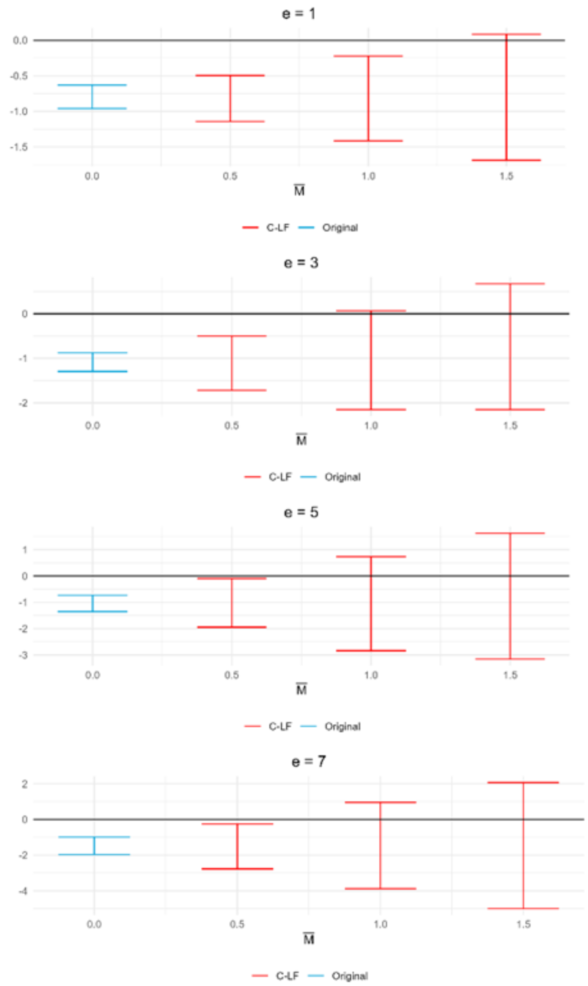
The "Honest DiD" approach uses these constraints to construct robust confidence intervals for the treatment effect. The idea is that even if the parallel trends assumption is not perfectly valid, we can still obtain valid inference on the treatment effect as long as the violations of parallel trends are limited by the imposed constraints.

Results in Figures C.1 to C.8

Figures A.1 to A.8 present the results of the "Honest DiD" sensitivity analysis for the four outcome variables considered in the study (age-grade distortion, IDEB scores, Portuguese scores, and mathematics scores). Each figure shows the robust confidence intervals for the treatment effect in different periods, under varying relative magnitude (RM(M)) and smoothness (SD(M)) constraints.

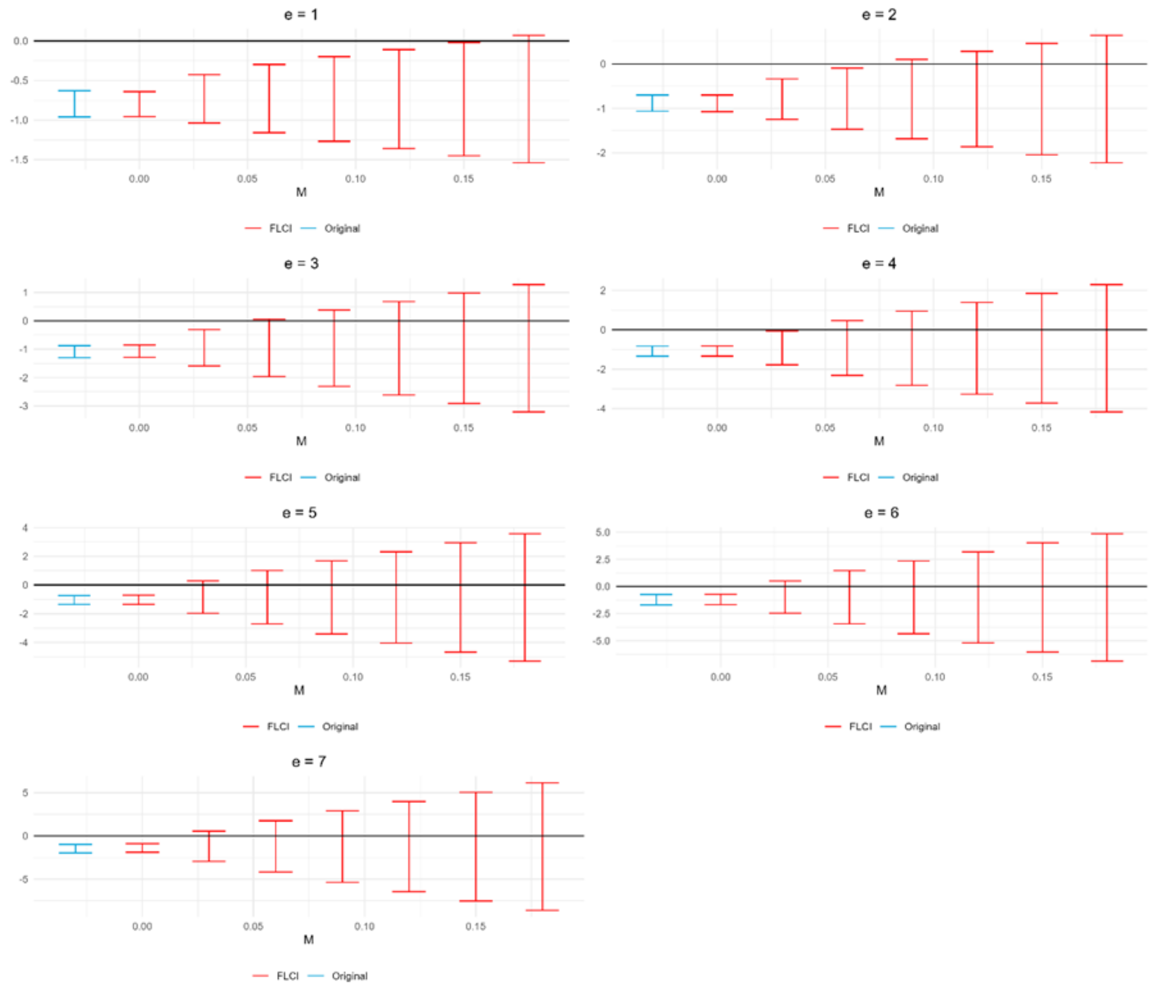
The results indicate that the estimated treatment effects are robust to different magnitudes of violations of the parallel trends assumption, particularly in the initial periods after the intervention. This suggests that even if the parallel trends assumption does not hold perfectly, the study's conclusions about the positive impact of militarization on student achievement in Goiás are robust to some degree of violation of this assumption.

Figure C.1 - Relative Magnitude Constraints - Age Distortion Series



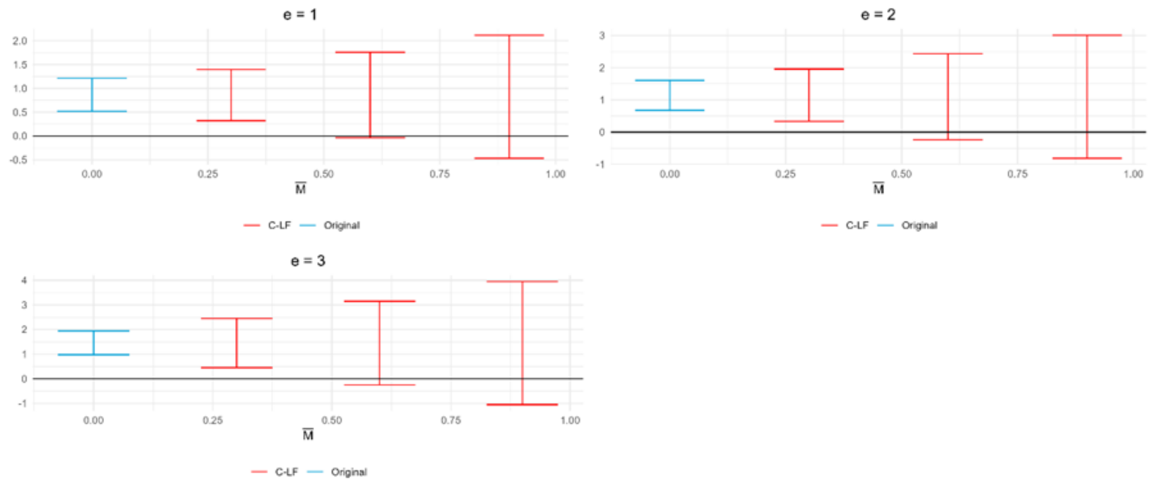
Note: Prepared by the authors.

Figure C.2 - Smoothness Constraints - Age Distortion Series



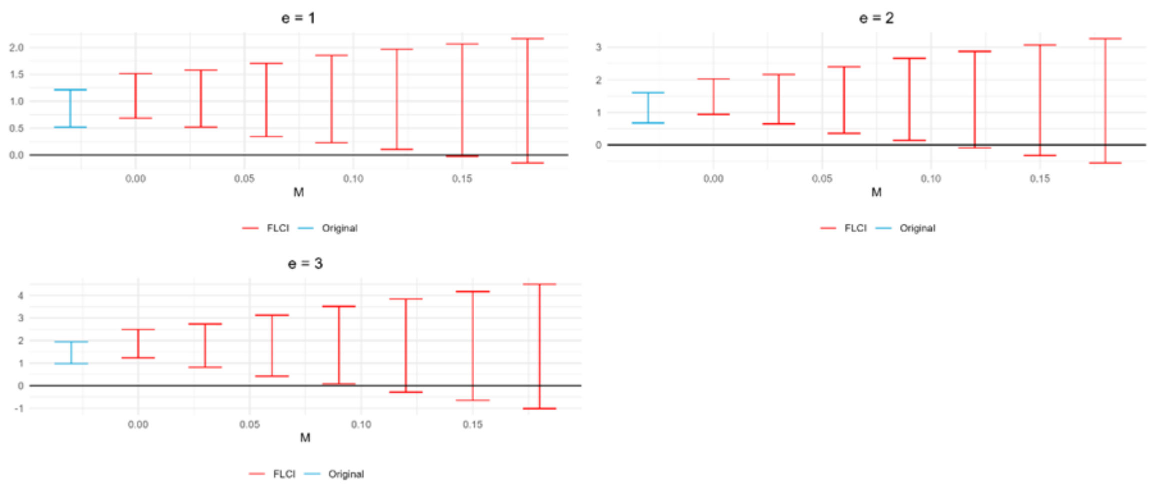
Note: Prepared by the authors.

Figure C.3 - Relative Magnitude Constraints - IDEB Note



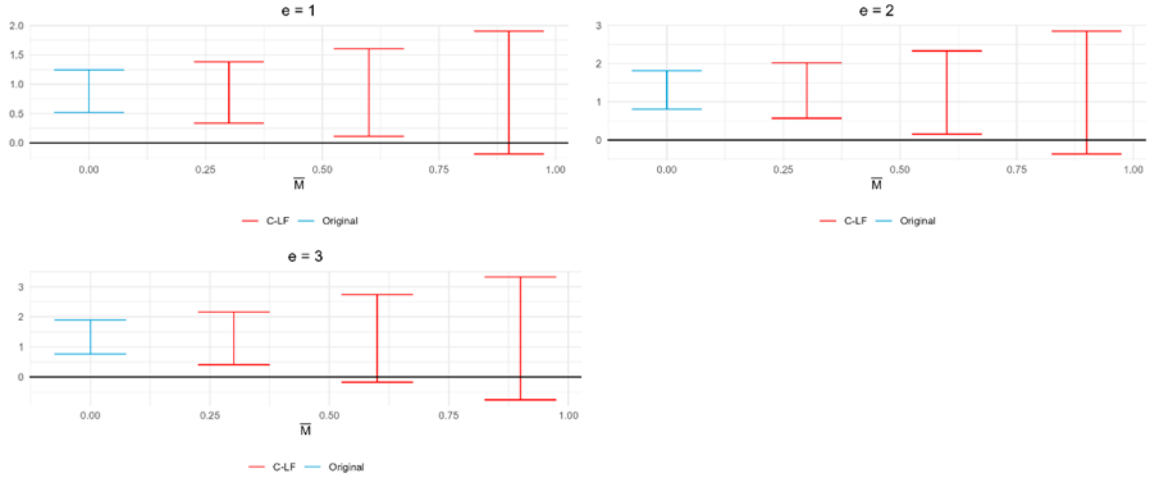
Note: Prepared by the authors.

Figure C.4 - Smoothness Constraints - IDEB No



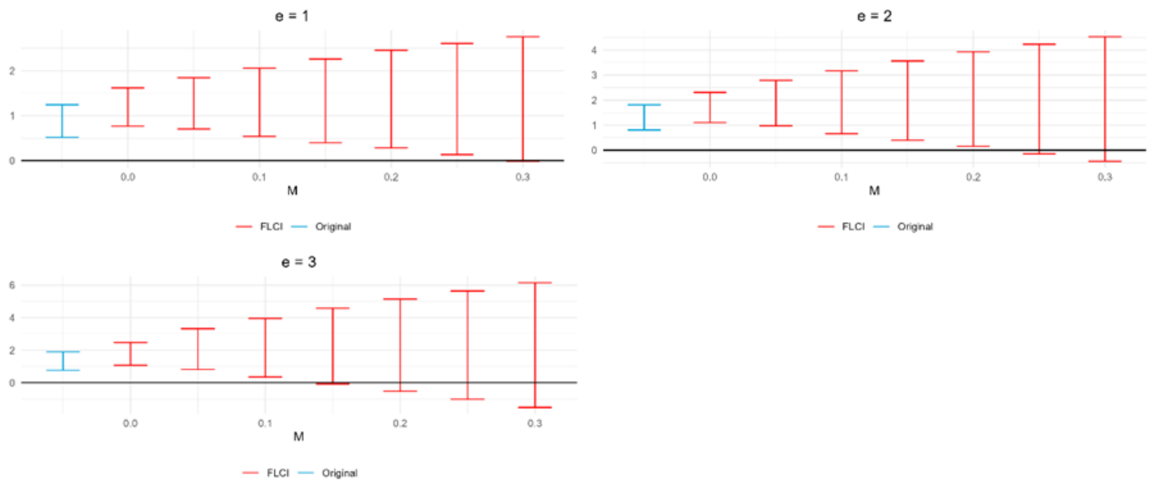
Note: Prepared by the authors.

Figure C.5 - Relative Magnitude Constraints - Mathematics Note



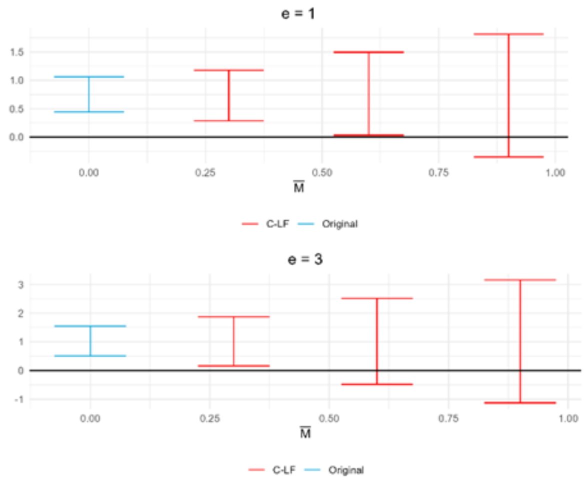
Note: Prepared by the authors.

Figure C.6 - Smoothness Constraints - Mathematics Note



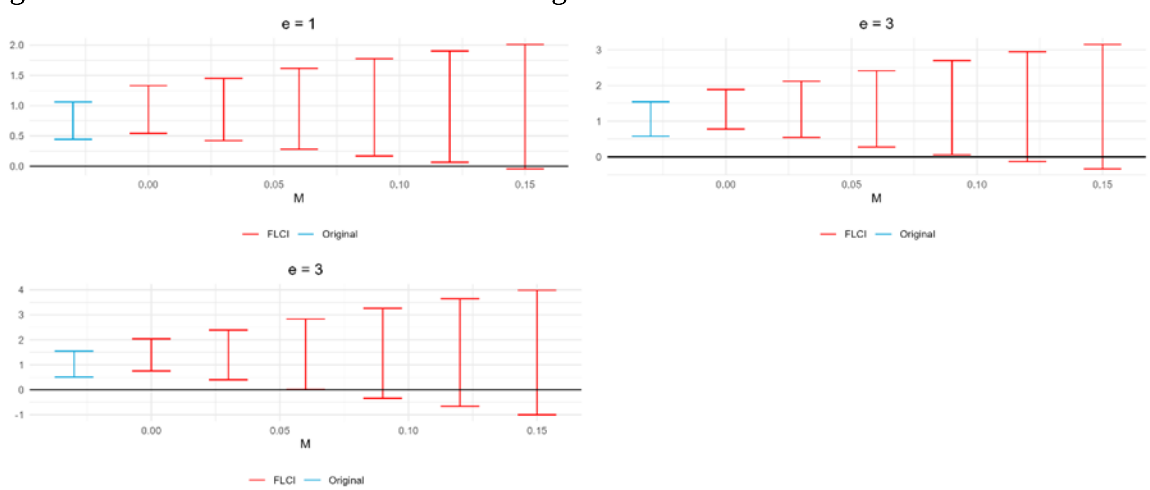
Note: Prepared by the authors.

Figure C.7 - Relative Magnitude Constraints - Portuguese



Note Prepared by the authors.

Figure C.8 - Smoothness Constraints - Portuguese Note



Note: Prepared by the authors.