Vivian Amorim

Essays on Public Primary Education in Brazil

Brasília/Brazil

December 2022

Vivian Amorim

Essays on Public Primary Education in Brazil

A thesis submitted to the *Universidade de Brasília* in partial fulfillment of the requirements of the Degree of Doctor of Philosophy in Economics

Universidade de Brasília

Faculdade de Administração, Economia e Contabilidade

Advisor: Michael Christian Lehmann

Brasília/Brazil

December 2022

Amorim, Vivian

Essays on Public Primary Education in Brazil / Vivian Amorim. Orientador: Michael Christian Lehmann. – Brasília, 2022.

151p.

Tese (Doutorado) - Universidade de Brasília, 2022.

1. Ensino Fundamental. 2. Diferença-em-diferenças. 3. DEA. 4. Avaliação não-experimental. 5. Gastos em Educação. I. Universidade de Brasília. II. Michael Christian Lehmann. III. Título.

Vivian Amorim

Essays on Public Primary Education in Brazil

A thesis submitted to the *Universidade de Brasília* in partial fulfillment of the requirements of the Degree of Doctor of Philosophy in Economics

Trabalho aprovado. Brasília, 9 de dezembro de 2022.

Michael Christian Lehmann

Orientador

Maria Eduarda Tannuri-Pianto

Membro titular vinculado

Paulo Augusto Meyer

Membro externo não vinculado à UnB

Caio Cordeiro Resende

Membro externo não vinculado à UnB

Rafael Terra

Membro Suplente

Brasília/Brazil

December 2022

Agradecimentos

Agradeço aos meus pais por terem me dado a oportunidade de estudar. A educação transforma vidas. Todas as oportunidades que tive são consequência da possibilidade de ter estudado. Agradeço às minhas irmãs por sempre terem acreditado em mim e por terem me dado tanto amor. Aos meus amigos de graduação e mestrado, obrigada pelo suporte acadêmico de sempre e por proporcionarem tantos momentos de descontração. A todos os meus demais amigos, obrigada por tudo, sempre. Ao Brian, obrigada pela companhia, paciência e por cuidar com tanto amor da minha pequena.

Agradeço ao meu orientador, Christian Lehmann, pela compreensão e pelo valioso suporte na condução deste trabalho. Ao Caio Piza por ter me dado a oportunidade de conciliar trabalho e estudo, além de todos as contribuições que proporcionaram enriquecer as análises empíricas. Eu nunca vou esquecer, não teria sido possível sem a sua ajuda. Ao Ildo Lautharte pelas suas contribuições. À Maria Eduarda Tannuri-Pianto, ao Paulo Augusto Meyer e ao Rafael Terra, membros da minha banca de qualificação, pelas comentários que proporcionaram mudanças significativas na tese.

"The content of a book holds the power of education and it is with this power that we can shape our future and change lives"

Malala Yousafzai

Abstract

The first chapter of this thesis investigates the impacts of the school closures adopted in São Paulo/Brazil amid the 2009 H1N1 outbreak. I find evidence that a three-week shutdown reduces in test scores equivalent to at least six weeks of schooling. The effects are more pronounced among the state-managed schools, where I estimate a decrease of 0.19 standard deviation in fifth graders' proficiency in Portuguese and a decrease of 0.26 standard deviation in students' proficiency in math. In locally-managed schools, the effects are restricted to math and are equivalent to a 0.18 standard deviation. The second chapter explores the impacts of *Acelera*, an intervention that has been implemented in Recife/Brazil since 2010 and focuses on primary education students who are at least one year older than the adequate age for their grade and who lag behind their peers. The program aims to increase learning levels and grade promotion, and decrease dropout and age-grade distortion. I do not find evidence that *Acelera* increases students' proficiency in reading and math. Nonetheless, my estimates suggest that the program increases grade promotion by 22.6% and decreases age-grade distortion by 17%. The heterogeneity analysis indicates that students with fewer years of age-grade distortion tend to benefit more from the intervention. The third chapter assesses the inefficiency of public primary education expenditures in Brazilian municipalities. I estimate that local authorities efficiently use between 72% to 83% of their resources. This means that by increasing their efficiency, for example adopting the best practices of the municipalities on the efficient frontier, there would be a fiscal space of at least 86 billion BRL, which is more than twice the 2022 *Bolsa Família* budget, the most import conditional cash transfer in the country. An amount that could be allocated to interventions to increase students' performance in a post-pandemic context where they are so much needed.

Keywords: 1. Primary Education. 2. Difference-in-differences. 3. DEA. 4. Non-experiemental evaluation. 5. Education Expenditures.

Resumo

O primeiro capítulo desta tese investiga os impactos do fechamento de escolas adotado em São Paulo/Brasil em meio ao surto de H1N1 de 2009. Meus resultados sugerem que o fechamento de três semanas leva a uma redução do aprendizado equivalente a pelo menos seis semanas de aula. Os efeitos são mais pronunciados entre as escolas estaduais, onde estimo uma diminuição de 0,19 desvio padrão na proficiência em português dos alunos da quinta série e uma diminuição de 0,26 desvio padrão na proficiência em matemática. Nas escolas municipais, os efeitos são restritos à matemática e equivalem a 0,18 desvio padrão. O segundo capítulo explora os impactos do *Acelera*, uma intervenção que vem sendo implementada em Recife/Brasil desde 2010, e foca nos alunos dos anos iniciais do Ensino Fundamental que apresentam pelo menos um ano de distorção idade-série e que têm desempenho inferior aos seus pares. O programa visa aumentar os níveis de aprendizado e a aprovação, e diminuir o abandono e a distorção idade-série. Não encontro evidências de que o *Acelera* aumenta a proficiência dos alunos em leitura e matemática. No entanto, minhas estimativas sugerem que o programa aumenta a aprovação em 22.6% e diminui a distorção idade-série em 17%. A análise de heterogeneidade indica que alunos com menos anos de distorção idade-série tendem a se beneficiar mais da intervenção. O terceiro capítulo avalia a ineficiência dos gastos públicos nos anos iniciais do Ensino Fundamental municípios brasileiros. Eu estimo que os governos municipais utilizam eficientemente entre 72% e 83% de seus recursos. Isso significa que com o aumento da eficiência, por exemplo por meio da adoção de algumas das melhores práticas dos municípios da fronteira de eficiência, haveria um espaço fiscal de pelo menos R$ 86 bilhões, mais que o dobro do orçamento do *Bolsa Família* para 2022, o maior programa de transferência de renda do país. Um montante que poderia ser destinado a intervenções para aumentar o desempenho dos alunos num contexto pós-pandemia em que são tão necessários.

Palavras-chave: 1. Ensino Fundamental. 2 Diferença-em-diferenças. 3. DEA. 4. Avaliação não-experimental. 5. Gastos em Educação.

Contents

1 Learning losses caused by school shutdowns: the impact of the H1N1 on Brazilian students in public primary schools **1**

1.1 Introduction . . . 1

1.2 Related Literature . . . 6

1.3 Background: the H1N1 outbreak and school shutdowns in São Paulo . . . 9

1.4 Data . . . 13

1.5 Empirical Strategy . . . 16

1.5.1 DiD on the sample of locally-managed schools . . . 19

1.5.2 Triple DiD on the sample of state and locally-managed schools . . . 21

1.6 Results . . . 24

1.6.1 Robustness . . . 29

1.6.2 Potential mechanisms . . . 31

1.7 Discussion and Conclusion . . . 38

2 *Teaching at the Right Level*: evidence from the Program *Acelera* in Brazil **41**

2.1 Introduction . . . 41

2.2 Related Literature . . . 47

2.3 The Intervention: *Acelera* . . . 50

2.4 Data . . . 54

2.5 Empirical Strategy . . . 62

2.5.1 Standard difference-in-differences . . . 62

2.5.2 Leads and lags . . . 64

2.6 Results . . . 65

2.7 Discussion and Conclusion 73

3 Estimating the efficiency of public primary education expenditures in Brazil 75

3.1 Introduction 75

3.2 Related Literature 78

3.3 Data 81

3.4 Empirical Strategy 85

3.5 Results 90

3.5.1 Input-oriented 90

3.5.2 Output-oriented 101

3.6 Discussion and Conclusion 104

A Appendix to Chapter 1 114

B Appendix to Chapter 2 135

C Appendix to Chapter 3 145

List of Figures

1.1 Students' proficiency in locally-managed schools, fifth-grade (2005-2009) 20

1.2 Impact of the school shutdowns by percentiles, fifth-grade (2007-2009) 28

2.1 Leads and lags estimates for grade promotion, pooled sample (2008-2014) 65

3.1 Distribution of per-pupil expenditure and socioeconomic background (2019) . . . 97

3.2 Boostrap bias and confidence intervals (2019) 99

3.3 Distribution of input-oriented efficiency scores, municipal and state governments (2019) 102

A.1 School shutdown policy, São Paulo (2009) 131

A.2 H1N1 confirmed cases per 100.000 inhabitants, São Paulo (2009) 132

A.3 Retention, dropout and grade-promotion, fifth-grade (2007-2009) 133

A.4 Students' proficiency in state and locally-managed schools, fifth-grade (2005-2009) 134

B.1 Distribution of students with at least one year of age-grade distortion, third to fifth grade (2010-2018) 140

B.2 Distribution of students per class, third to fifth grade (2010-2018) 141

B.3 Retention, dropout and grade-promotion of Regular and *Acelera* students, third to fifth grade (2010-2014) 142

B.4 Performance in reading and math, third and fifth grade (2010-2018) 143

B.5 Propensity score matching on the sample of schools offering *Acelera*, first to fifth grade (2010-2014) 144

C.1 Distribution of input-oriented efficiency scores, with and without non-discretionary inputs (2019) 150

C.2 Distribution of input-oriented efficiency scores (2013-2019) 151

List of Tables

1.1 State and locally-managed schools, São Paulo (2009) 12

1.2 Sample of the study, São Paulo (2005-2009) 18

1.3 Impact of the school shutdowns on students' learning, fifth-grade (2007-2009) . . 25

1.4 Principals heterogeneity on the impact of the school shutdowns on students' learning, fifth-grade (2007-2009) 34

1.5 Teachers' heterogeneity on the impact of the school shutdowns on students' learning, fifth-grade (2007-2009) 36

2.1 Theory of Change of *Acelera* 53

2.2 Sample of the study, first to fifth grade (2010-2018) 58

2.3 Sample of fourth graders, (2011-2017) 61

2.4 Impact of *Acelera* on grade promotion, dropout, and age-grade distortion, first-fifth grade (2008-2014) 68

2.5 Heterogeneity of the impact of *Acelera* on grade promotion, dropout, and age-grade distortion, first-fifth grade (2008-2014) 70

2.6 Impact of *Acelera* on students' performance in reading and math, fourth grade (2008-2018) 72

3.1 DMUs, output, and inputs of the DEA 84

3.2 Input-oriented efficiency scores (2019) 93

3.3 Example of benchmark DMUs (2019) 95

3.4 Output-oriented efficiency scores (2019) 103

A.1 Balance test: municipalities that extended the winter break versus municipalities that did not (2009) 114

A.2 Linear Regression of 2007 school characteristics on the decision to extend the winter break in 2009 115

A.3 Fifth-graders' characteristics, locally-managed schools (2009) 116

A.4 Fifth-graders' characteristics, state and locally-managed schools in $G = 1$ (2009) 117

A.5 Fifth-graders' characteristics, state and locally-managed schools in $G = 0$ (2009) 118

A.6 Fifth-graders' characteristics, state-managed schools (2009) 119

A.7 Teachers' and principals' characteristics, locally-managed schools (2009) 120

A.8 Teachers' and principals' characteristics, state and locally-managed schools in $G = 1$ (2009) 121

A.9 Teachers' and principals' characteristics, state and locally-managed schools in $G = 0$ (2009) 122

A.10 Teachers' and principals' characteristics, state-managed schools (2009) 123

A.11 Placebo test for students' performance in treatment and comparison groups (2005-2007) 124

A.12 Placebo test for grade-promotion, retention, and dropout (2007-2008) 125

A.13 Impact of the school shutdowns on students' learning (with placebo), fifth-grade (2005-2009) 126

A.14 Impact of school shutdowns on grade-promotion, retention and dropout, fifth-grade 127

A.15 Impact of the school shutdowns on students' learning (with and without schools with managerial practices intervention), fifth-grade (2007-2009) 128

A.16 Impact of the school shutdowns on students' learning (with and without controls), fifth-grade (2007-2009) 129

A.17 Average number of state and locally-managed schools included in the triple DiD (2007-2009) 130

B.1 Educational indicators, first to fifth grade (2007-2019) 136

B.2 Percentage of students in the EMPREL dataset whose proficiency in reading and math is available (2010-2018) 137

B.3 Balance test between Regular and *Acelera* schools (2010-2018) 138

B.4 Balance test between *Acelera* participants that have and that do not have proficiency in fifth grade available (2010-2018) 139

C.1 Price deflator by geographical area (2008) 146

C.2 Descriptive statistics of the inputs and outputs according to municipality size (2015-2019) 147

C.3 Sample-size of the Brazilian municipalities included in the DEA (2019) 148

C.4 Number of benchmark DMUs by state (2019) 149

C.5 Input-oriented efficiency scores (2013-2019) 149

1 Learning losses caused by school shutdowns: the impact of the H1N1 on Brazilian students in public primary schools

1.1 Introduction

School closures are one of the most drastic measures to contain the spread of infectious diseases. Existing evidence shows that school shutdowns, even for a short period of time, can have large negative effects on learning (Andrabi et al. (2020), Marcotte and Hemelt (2008), Donnelly and Patrinos (2021)). In 2009, the alarming rapidity with which the H1N1 virus, commonly known as swine flu, was spreading led some countries to adopt this strategy, such as Brazil, Mexico, and the United States. In the state of São Paulo, the largest public school system in Brazil, more than half of the primary, lower, and upper secondary schools were closed between two to three weeks, affecting more than 5.5 million students (70% of the students in the state). I leverage this natural experiment to estimate the impact of school closures on students' learning.[1]

The state of São Paulo has 645 municipalities, and the state and municipal governments share the responsibilities for the provision of public education. In each one of these municipalities, there is at least one school managed by the state authority (state-managed school) and one school managed by the respective local authority (locally-managed school).[2] The main responsibilities of both governments are to hire teachers, provide textbooks, appoint principals, finance school infrastructure, and determine the length of the school day, as well as the beginning and the end of the school year in the schools under their responsibility. In July of every school year, schoolchildren have their winter break. For primary education, the focus of this chapter, the state, and municipal governments manage schools in 206 and 642 of the 645 municipalities, respectively.[3]

During the H1N1 outbreak in 2009, the Department of Health of the state of São Paulo recommended the postponement of children's winter break. The state government enforced this policy on all their state-managed schools. The municipal governments had the autonomy

[1] I focus on the state of São Paulo since the state has the largest public network in the country and where I could find, by checking local newspapers, the name of the municipalities whose local authorities opted to close the schools under their management.

[2] I use the terms municipal government and local authority interchangeably.

[3] Census of Education, 2009.

to decide whether or not to follow the same guideline in the schools under their management. In the end, thirteen of the 642 local authorities that provide primary education extended children's winter break in their schools. The remaining municipal governments followed the calendar as previously planned, providing a group from which the closed schools can be compared.[4] Overall, more than three thousand state and locally-managed schools offering primary education were closed, and almost five thousand locally-managed schools remained open (Table 1.1).

Among the factors that might have played an important role in the local authorities' decision to impose school shutdowns are population size, number of confirmed H1N1 cases per thousand inhabitants, the capacity of the health system, number of students per school and per class, and power of teacher's union (a fact that can be associated with the percentage of teachers with tenure). Indeed, six of the thirteen municipalities in which the local authority decided to close its locally-managed schools are among the nine more populous in the state of São Paulo. On average, these municipalities registered more confirmed cases of H1N1 per 100 thousand inhabitants between April and July (10×2 cases, Figure A.2). Also, the percentage of teachers with tenure in their locally-managed schools is 20 percentage points higher (Table A.1).

In this chapter, I test the hypothesis that school closures decrease primary education students' proficiency in reading and math. I rely on the results of the national proficiency assessment, *Prova Brasil*, applied to fifth-grade students. The exam collects information on students' proficiency in reading and math, socioeconomic indicators, and teachers' and principals' characteristics. I complement the analysis with additional information from the Census of Education, a dataset that contains several variables such as schools' amenities, school hours, length of the school year, and pupil-teacher ratio. The regressions are run at the school level and I compare distinct cohorts of fifth-graders between 2007, before the shutdowns, and 2009, the year the schools were closed.[5]

[4] The municipalities whose local authorities decided to extend the children's winter break are presented in Figure A.1.

[5] The data at the student level would be more adequate as I would have more observations and my estimates could be more precise. However, due to time constraints, I conduct the analysis at the school level. Setting up the dataset at the school level was already an extensive task that took more than one year since I work with four distinct microdata from the National Institute of Education and Research. I use the years from 2005 to 2009, requiring the harmonization of all the variables over the years (making sure that the variables have the same definition over the years and correcting inconsistencies). My code can be assessed in this link: https://github.com/worldbank/h1n1-school-closures-sp-2009.

I first estimate the impact of school closures on locally-managed schools by exploring the policy variation between municipalities under a difference-in-differences (DiD) design. I compare the locally-managed schools of the municipalities whose local authorities extended children's winter break with the locally-managed schools of the municipalities whose local authorities followed the school calendar as previously planned. In order to test for parallel trends before the shutdowns, the analysis is restricted to locally-managed schools that have proficiency data available since 2005.[6] This is the case for ten of the thirteen municipalities that postponed children's return to school and for 469 municipalities that did not. The DiD sample has 795 locally-managed schools that were closed and 2,568 that remained open (Table 1.2). To account for the factors that might have influenced the mayors' decision to impose school shutdowns and that are correlated with the students' performance, I include municipalities' fixed effects, and several controls that summarize students', teachers', principals', and schools' characteristics.

Under the DiD approach, one may wonder whether the influence of unobserved time-varying factors might affect differently the municipalities where the local authorities postponed children's return to school and the municipalities where the local authorities did not. If there are unobservable time-varying factors correlated with both the mayor's decision to extend the winter break and students' proficiency, the estimates would be biased. For example, the higher incidence of H1N1 cases in the municipalities that postponed the return to school could result in more students taking care of relatives, which would lead to less time to study. Also, student and teacher absenteeism and psychological distress could have increased at a higher pace in municipalities more affected by the pandemic.

To overcome this identification threat, I leverage a within-municipality variation created by the school shutdown policy that allows me to estimate the impact of the shutdowns on state-managed schools. For the same municipality, I am able to compare its state and locally-managed schools where, therefore, children are affected by the same shocks at the municipality level. I then restrict the sample to the municipalities where there is at least one state-managed school and one locally-managed school. This is the case for 112 municipalities whose local authorities did not extend the winter break (a group that I classify as $G = 1$), and for ten out of the thirteen municipalities whose local authorities postponed the children's return to schools ($G = 0$). I

[6] Not all the public schools are included in *Prova Brasil*, mostly due to their size as the exam is restricted to classes that have at least 20 students in the fifth grade.

employ a triple difference-in-differences approach in order to account for the differences in the learning trajectories of state and locally managed schools.

In the first DiD of the triple difference design, for schools located in $G = 1$, I compare 929 state-managed schools that were closed with 1,334 locally-managed schools that remained open to estimate the effects of the school shutdowns and the proficiency gap between state and locally-managed schools. In the second DiD, for schools located in $G = 0$, I compare 868 state-managed schools with 759 locally-managed schools, all closed, to estimate the learning gap between state and locally-managed schools. Finally, the third difference is the result of the subtraction of the first DiD and the second DiD and it is the estimate of the effect of the shutdowns on the state-managed schools. To account for the differences between state and locally-managed schools inside the same municipality, I add several controls that summarize students', teachers', principals', and schools' characteristics. Also, I control for school-fixed effects to account for non-observable time-invariant factors that might be correlated with the treatment and students' performance.

The results show that extending the winter break by two to three weeks led to a reduction in math scores of at least 0.18 of a standard deviation in locally-managed schools and 0.26 in state-managed ones, equivalent to at least six weeks of learning loss. I find evidence that the impacts were slightly stronger in schools in the bottom deciles of the math test score distribution, suggesting that the impacts were more pronounced among students lagging behind in mathematics. For Portuguese, the effects are restricted to schools managed by the state authority and reach at least 0.19 of a standard deviation, suggesting that the effects can vary by the school's level of administration.

I complement the analysis by exploring three potential mechanisms: principals' managerial skills, teacher absenteeism, and a shorter time frame to cover school curricula in closed schools. First, for the state-managed schools, I find suggestive evidence that the higher the teachers' perception of principals' skills, the more the negative impacts of the shutdowns are offset. Second, in state-managed schools where teacher absenteeism is seen as a big issue, the learning loss is at least 60% higher compared to state-managed schools where absenteeism is not a concern. Also, the data do not indicate that state-managed schools extended the length of the school year to compensate for the period of shutdowns. For state and locally-managed schools, it is likely that the rush to cover the school curriculum made it challenging for students to keep pace with it. Finally, for locally-managed schools, the shorter length of an average school day also seems to have been a concern.

This chapter presents several contributions to the literature. First, to my knowledge, this is the first work that assesses the effects of school closures by schools' level of administration. My results suggest that locally-managed schools tend to cope better with the shutdowns. This result is in line with the decentralization literature that points out that local governments, for being closer to the population, can better identify societies' needs. Second, I investigate how some schools' characteristics are able to offset the adverse effects of school closures, providing direct policy recommendations. Third, I take advantage of a natural experiment to estimate the effects of school closures using administrative datasets available before and after the shutdowns. Fourth, since most of the available evidence on the effects of school shutdowns is for developed countries (Donnelly and Patrinos (2021)), my work provides treatment effect estimates amid a pandemic episode in a developing country. Also, Brazil is the developing country most hit by the recent pandemic outbreak, Covid-19.[7] My estimates show the sizeable negative effects of school closures, even when it is for a short period of time. The results can shed light on the challenges the country faces after most of its public schools were closed for more than one year due to Covid-19.

In the process of returning to school, some of the main policies recommended to mitigate the adverse effects include: administering proficiency tests to identify the content that should be prioritized and the most vulnerable students who will need special attention; lengthening the school day; shortening planned holidays for December and January; promoting emotional support and campaigns to raise awareness about returning to school (via TV or digital media); improving distance learning platforms to complement classroom learning; and developing strategies to reduce teachers' absenteeism. Based on a Literature Review, Knoster (2016) recommends the following strategies to reduce teachers' absenteeism: rewarding teacher attendance and performance and encouraging collegial relationships among teachers and leaders. Providing emotional support for students who faced trauma is critical. There is a significant body of evidence suggesting that counseling can improve student grades, attendance, behavior, and graduation (American School Counselor Association (2015)).

[7] In terms of the number of deaths and H1N1 cases by 100.000 inhabitants.

Awareness campaigns might also be effective in avoiding student dropouts, especially among girls and other vulnerable groups (Rogers and Sabarwal (2020)).

Apart from this introduction, this chapter is organized as follows: Section 1.2 presents the related literature. Section 1.3 summarizes the H1N1 outbreak. Section 1.4 introduces the data available to perform the analysis. Section 1.5 discusses the empirical strategy. Section 1.6 presents the main findings. I then conclude with a discussion and policy implications in Section 1.7.

1.2 Related Literature

The strategy to impose school shutdowns during pandemics dates back to 1916 when several cities in the United States decided to impose this policy during the polio outbreak. This strategy affects learning levels that might lead to long-lasting consequences. Younger children have their learning disrupted in a critical period of development. Literacy and numeracy instruction in early grades constitutes the foundation that future learning will take place (Rogers and Sabarwal (2020)). Furthermore, since knowledge is cumulative, if children fail to achieve the right set of skills in elementary education, they will struggle to learn in later grades (Crouch and Gove (2011)). Meanwhile, older children may opt to drop out of school to join the labor market or even due to teenage pregnancies and early marriages.

In 1916, several cities in the USA imposed school shutdowns for up to one month as an attempt to control the spread of the Polio pandemic. Meyers and Thomasson (2017) explores the fact that the cities more affected by the outbreak were the ones to impose school closures. The authors then use morbidity rate as a proxy for school shutdowns. They find evidence that a 1% increase in the morbidity rate resulted in people aged 14 to 17 during the outbreak having a 6% reduction in their average educational attainment 22 years after it. The authors claim that children of legal working age might have decided to join the labor market and not return to school after the reopening.[8]

The school shutdowns tend to widen gender gaps because of girls' increased dropout rates. As a result of being out of a protective environment, girls are more vulnerable to sexual abuse.

[8] By that time, the legal working age was above 13 in most states of the USA.

During the Ebola outbreak, higher rates of transactional sex, and early and forced marriages were reported as families struggled to cover basic needs. Teenage pregnancy increased by 65% in Sierra Leone and girls were 16 percentage points less likely to be in school after the reopening (United Nations Development Programme (2015), Rogers and Sabarwal (2020)). Since the Covid-19 outbreak in March 2020, more than 20 thousand girls aged 10 to 14 have become mothers in Kenya, a country where 2 out of 5 teenagers are either pregnant or mothers already.[9] All these factors decrease the likelihood that affected girls will return to school.

Another starting point to estimate the consequences of school shutdowns is to explore the effects of other situations that lead children to be away from school. Marcotte and Hemelt (2008) assess the impacts of unscheduled school closings due to more severe winters in Maryland/USA. Using data from 1994 to 2005, the authors find evidence that each school day lost reduces the percentage of third-graders who perform satisfactorily in reading and math by 0.5%. In a typical winter with an average of 5 days of unscheduled closures, nearly 3% fewer students pass reading and math tests. Their results also suggest that the longer the period away from school, the more seriously disruptive the effect on math performance. It is estimated that in years with 3 to 5 days of closures, each day lost causes 0.25% fewer students to achieve a satisfactory proficiency, while in years with 8 to 10 days, the percentage is 0.33%, and it reaches 0.5% when students lose more than 12 days.

Andrabi et al. (2020) study the effects of the 2005 earthquake in Pakistan on human capital accumulation. The authors find evidence that four years after the incident, there was a full recovery of a large number of household and adult outcomes, which is likely to be explained by the massive aid from the government. However, there was a significant shortfall in the learning levels of children aged 3 to 11 during the earthquake. Children who lived at a maximum distance of 20 km from the fault line, where schools remained closed for an average of 14 weeks, had test scores of 0.4 of a standard deviation lower than children who lived farther from it. This result is equivalent to two years of schooling. The earthquake not only increased inequality across villages but also sharpened the differences within them, as the losses were all felt by children whose mothers do not have primary education.

Belot and Webbink (2010) investigates the long-term effects of a 5-month school interruption

[9] Results available in this link: https://citizentv.co.ke/news/president-uhuru-on-teenage-pregnancies-there-will-be-hell-to-pay-for-chiefs-perpetrators-337610/.

caused by the 1990 teacher strike in Belgium. The results indicate that the educational attainment of the affected cohort lags behind by 0.7 years, which is likely due to the increase in repetition rates. Since there was no official change in the school calendar after the strikes, students entered the next level less prepared and were probably not able to catch up with the missing content, causing higher repetition. Besides that, the probability of having a university diploma fell by 2% in the cohort affected by the strikes.

School closures are likely to increase the learning gap between students from high and low socioeconomic backgrounds. The literature on learning loss that happens during summer breaks in the United States can shed light on this subject. Research has shown that students' skills and knowledge often deteriorate during the 3-month summer vacations, with low-income students facing more substantial losses (McCombs (2011)).

Alexander et al. (2007) estimates that two-thirds of the achievement gap of ninth graders from high and low socioeconomic backgrounds in Baltimore/USA could be attributed to summer learning loss in the first five years of schooling. The learning gap might help to explain the higher dropout rates of vulnerable students, the lower percentage of them that follow a college-preparatory high school program, and that go to college.

A meta-analysis conducted by Cooper et al. (1996) indicates that students from grades 1-9 lost the equivalent of one month of instruction during the three months of summer break in the USA. The detrimental effects are higher for math than for reading. While for math, the impacts are harmful regardless of the income level, the decrease in reading skills is concentrated among low-income students.

Children tend to forget more math than reading as the majority of their exposure to math happens in school from teachers. On the other hand, reading skills are more affected by factors outside school as families are more likely to promote and practice literacy skills at home. Exposure to reading can vary with socioeconomic background, with low-income students falling behind during school breaks and high-income students making gains. Different availability of opportunities to practice reading can explain these results, as wealthier children have more access to books and are more prone to participate in summer activities that require literacy skills (McCombs (2011)).

Traumatic experiences affect the potential to learn. A pandemic that involves death, loss, insecurity, social isolation, and increased exposure to domestic violence can affect children's well-being and learning. The Global Education Monitoring Report (2019) points to the disruptive effects on learning as physical and emotional, involving anxiety, fear, lack of emotional control, and sadness. There are also effects on children's cognitive skills, expressed via difficulty in paying attention, inability to process information, and memory problems. Besides worsening the interaction with other students, these factors might affect the student-teacher relationship.

Based on the results of the previous research, I test the hypothesis that school closures decrease primary education students' proficiency in reading and math. I investigate how the impact of the shutdowns varies according to schools' level of administration, that is, whether state or locally-managed schools are better able to cope with it. I further enrich the analysis by exploring some factors, such as principals' managerial skills and teachers' absenteeism, that can either offset or worsen the impacts of school closures on students' performance.

1.3 Background: the H1N1 outbreak and school shutdowns in São Paulo

In June 2009, a new influenza outbreak, the H1N1, was declared a pandemic by the World Health Organization (WHO).[10] From April to December 2009, Brazil confirmed 54,171 cases and 2,196 deaths from H1N1.[11] However, the number is likely to be much higher as many people with flu symptoms do not seek help, not all the ones that look for health assistance are tested, and the under-reporting of hospitalizations and deaths are publicly known.

In July of 2009, children were on the winter break of the school year. Amid the increase in the number of cases, the Departments of Health of several Brazilian states, such as São Paulo, Minas Gerais, Rio de Janeiro, Paraná, Rio Grande do Sul and Distrito Federal, recommended that schools extended the winter break to avoid the spread of the disease, as this policy is an important component of a community's social distancing efforts (Adda (2016)).[12] Children are

[10] By May 2010, 214 countries had reported cases and an estimated death toll of more than 200,000 people (The World Health Organization (2010), Dawood et al. (2012)).

[11] DATASUS, 2009. Available in this link: http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinannet/cnv/influbr.def.

[12] See the list of the states in this link: https://www1.folha.uol.com.br/cotidiano/2009/08/602634-escolas-e-universidades-adiam-volta-as-aulas-devido-a-gripe-suina-veja-lista.shtml. Check the official statement of the state government of São Paulo in this link: https://www.saopaulo.sp.gov.br/ultimas-noticias/nota-oficial-da-secretaria-da-saude-sobre-retorno-as-aulas/.

less likely to adopt behavioral changes, such as washing hands and reducing physical touch, more prone to sustain person-to-person contact for prolonged times, and in some cases, more susceptible to infectious diseases, potentially acting as a vector of transmission (Klaiman et al. (2011)).[13]

The state government postponed the return to school for two weeks in all the state-managed schools across all 645 municipalities. The municipal governments were free to decide whether to follow the same guidelines in the schools under their management. By checking local newspapers, I find that thirteen local authorities followed the guidance of the Department of Health: São Paulo, the state capital, Campinas, Diadema, Embu das Artes, Indaiatuba, Mairiporã, Osasco, São Bernardo do Campo, Santo André, São Caetano do Sul, Sumaré, Ribeirão Preto and Taboão da Serra (Figure A.1). These municipalities postponed the winter break of their locally-managed schools for two to three weeks. The remaining municipal governments in the state followed the calendar as previously planned, providing a group from which the closed schools can be compared.

Table 1.1 shows that, on the one hand, 12,957 state and locally-managed schools were closed, which represent more than half of the public schools in the state of São Paulo, affecting more than 5.5 million students (70% of the students in the state).[14] On the other hand, 12,192 locally-managed schools with almost 2.5 million students remained open.

Since the state government extended the winter break of all the state-managed schools across the 645 municipalities, but only thirteen municipal governments decided to adopt the same measure in the schools under their management, it is important to investigate what factors influenced their mayor's decision. Among the factors that might have played an important role are population size, the number of confirmed H1N1 cases per thousand inhabitants, the capacity of the health system, the number of students per school and per class, and the power of teacher's union (a fact that can be associated with the percentage of teachers with tenure). Indeed, six out of the thirteen municipalities in which the municipal government decided to

[13]Another countries, such as the USA and Mexico, also adopted social distancing measures to slow the spread of the disease. The guidelines included staying home when ill unless to seek medical care, avoiding large gatherings, telecommuting, and closing schools. In Mexico City, the shutdowns affected 7 million students from kindergartners to college (Lacey and McNeil Jr (2009)). In the USA, the Centers for Disease Control and Prevention (CDC) recommended school closures for 14 days if H1N1 was identified among the students. The country closed 726 primary, lower, and upper secondary schools, affecting 368,282 students (Klaiman et al. (2011)).

[14]Census of Education, 2009.

close the locally-managed schools are among the nine more populous in the state of São Paulo. All the thirteen ones are between one-fifth more populous. On average, these municipalities registered more confirmed cases of H1N1 per 100 thousand inhabitants between April and July (10×2 cases, Figure A.2). Also, the percentage of teachers with tenure in their locally-managed schools is 20 percentage points higher (Table A.1).

Table 1.1: State and locally-managed schools, São Paulo (2009)

	Students					Schools				
	Total	Pre-K	Primary Education	Lower Secondary	Upper Secondary	Total	Pre-K	Primary Education	Lower Secondary	Upper Secondary
Schools shutdown, extension of the winter-break										
State-managed schools located in all the 645 municipalities in the state of São Paulo	4,338,887	1,193	852,116	1,870,373	1,615,205	9,787	39	2,144	3,833	3,771
Locally-managed schools located in 13 out of the 645 municipalities in the state of São Paulo	1,202,386	454,499	454,703	281,392	11,792	3,170	1,579	972	596	23
Total	5,541,273	455,692	1,306,819	2,151,765	1,626,997	12,957	1,618	3,116	4,429	3,794
No school shutdown, no extension of the winter break										
Locally-managed located in 632 out of the 645 municipalities in the state of São Paulo	2,458,858	727,745	1,380,053	326,645	24,415	12,192	6,204	4,956	947	85

% of students affected by the shutdowns	% of schools affected by the shutdowns
69.3%	51.5%

Notes: Pre-k: Kindergarten and pre-school. Primary Education: first to fifth grades. Lower secondary: sixth to ninth grades. Upper secondary education: tenth to twelfth grades (high school). The locally-managed schools: schools managed by the local authorities of the 645 municipalities in São Paulo. The state-managed schools: schools managed by the state government of São Paulo.
Source: Census of Education, 2009.

It is also important to understand the main differences between the locally-managed schools of the municipalities whose local authorities extended the winter break and the locally-managed schools of the municipalities whose local authorities followed the school calendar as previously planned. I then run a regression of a dummy indicating whether the school s in municipality m was closed in 2009 on 2007 data of GDP per capita of the municipality, municipality population size, number of confirmed cases of H1N1 per 100 thousand inhabitants, and students', teachers' and principals' characteristics. I find that the main differences between them are that the locally-managed schools of the municipalities whose local authorities extended the winter break had shorter school days (4.8 hours in affected schools $\times$ 5 hours in the unaffected ones), a higher percentage of principals that see teachers' absenteeism as a big issue (26.7% $\times$ 7.7%), as well as students' absenteeism (8.7% $\times$ 4.2%), a higher percentage of teachers with tenure (64.7% versus 49%), and lower levels of principals' managerial skills according to teachers' perspective (0.72 $\times$ 0.77) (Table A.2).[15]

1.4 Data

In Brazil, the public education system is decentralized and the 26 states, the Federal district, and the 5,570 municipalities share the responsibilities for the provision of education. According to the 1988 Constitution, the municipal governments should give priority to early childhood and primary (grades 1 to 5) and lower secondary education (grades 6 to 9), and the state authorities to primary, lower, and upper secondary education (grades 10 to 12). Hence, the schools located in all the Brazilian municipalities can be managed either by the state government or can be under the management of the municipal government. The first group of schools is state-managed, and

[15] All the descriptive statistics presented in this paragraph are relative to 2007 and based on *Prova Brasil* questionnaires that are filled out by students, teachers, and principals. Tables upon request. The index of principal managerial skills ranges from 0 to 1. It is calculated based on teachers' answers of how frequently they believe that the principal pays attention to students' learning, administrative norms, and school maintenance; motivates the teachers and encourages new ideas, and takes into consideration teachers' inputs; and whether teachers trust the principal and can participate of the decisions related to their work. All these variables have four possible answers: never (value 0), sometimes (0.33), often (0.66), and always (1). The principal managerial skills index at the teacher level is an average of these answers, and the index at the school level is an average of teachers' answers. Student motivation is an index ranging from 0 to 1. It is calculated based on the teachers' answers on whether students' learning deficit is caused by low student motivation or bad behavior in class. These variables have two possible answers: yes (0) and no (1). The variable students' absenteeism is also part of the index and has three possible answers: a moderate/big issue (0), a small issue (0.5), and not a problem (1). The student motivation index is an average of these answers. Teacher motivation is an index ranging from 0 to 1. It is calculated based on students' answers to how frequently the teacher corrects their Portuguese and math homework. The variable has three possible answers: always (1), sometimes (0.5), and never (0). Teachers' motivation at the school level is an average of students' answers.

the second group of schools is locally-managed. In each one of the 645 municipalities in the state of São Paulo, there is at least one locally-managed school and one state-managed school offering either preschool, primary, lower, or upper secondary education.

Article 211 - 1988 Brazilian Constitution
§ 2o Os Municípios atuarão prioritariamente no ensino fundamental e na educação infantil. "Municipalities will give priority to providing early childhood, primary, and lower secondary education."
§ 3o Os Estados e o Distrito Federal atuarão prioritariamente no ensino fundamental e médio. "States will give priority to providing primary, lower, and upper secondary education."

The Brazilian legislation for primary, lower, and upper secondary education determines a school year with a minimum of 200 days and 800 hours of instruction time. The state-managed schools are monitored by the State Department of Education, under the state government administration, and locally-managed schools are monitored by the Municipal Department of Education, under the municipal government administration. The respective Departments of Education are in charge of educational policies implemented at the school level, hiring teachers, providing textbooks, appointing principals, financing school infrastructure, and determining the length of the school day, and school breaks, as well as the beginning and the end of the school year in the schools under their responsibility. Usually, the school year goes from February to November, summer break occurs in December and January, and winter breaks cover two weeks in July.

To assess the impact of the 2009 school shutdowns on primary education students' proficiency in Portuguese and math, I use data from the Census of Education, the Brazilian Institute of Geography and Statistics (IBGE), and from *Prova Brasil*, which is the national proficiency exam to assess students' learning levels.[16]

Since 1995, all private and public schools offering primary, lower, and upper secondary education participate in the annual Census of Education.[17] The Census is implemented by the National Institute of Educational Studies and Research (INEP), a research agency under the Brazilian Ministry of Education.[18] The Census collects information on (i) school facilities,

[16]IBGE stands for *Instituto Brasileiro de Geografia e Estatística*. IBGE has information on GDP per capita and the population at the municipality level.

[17]Schools offering early childhood education also participate.

[18]INEP stands for *Instituto Nacional de Estudos e Pesquisas Educacionais*.

such as libraries, sports courts, and science and computer labs; (ii) school infrastructure, such as filtered water, electricity, and internet access; (iii) social services, for example, school transportation and provision of meals; (iv) students, such as sex, the color of the skin, age, physical disabilities or mental illness, grade level, instruction time per day, class-size, subjects they are enrolled in, grade promotion, repetition and dropout rates; and (v) teachers, such as educational attainment, age, physical disabilities, subjects taught, and classes they are in charge of.

Every two years, the INEP applies a national exam, *Prova Brasil*, to assess students' proficiency in Portuguese and math. Since 2005, the test is applied to fifth and ninth-graders in all public schools.[19] *Prova Brasil* is one of the proficiency tests within the scope of the Education Assessment System (SAEB).[20] Students take the test at the end of the school year (between October and November). In 2007, children took the test between November 5 and 20; and in 2009 between October 19 and 31 (approximately, two months after the school shutdowns). For fifth-graders, proficiency in Portuguese has a scale ranging from 0 to 325, and math has a scale ranging from 0 to 350 (SAEB scale). The students' proficiency can be classified as insufficient, basic, or advanced.

In addition to answering Portuguese and math questions, students fill out a socioeconomic questionnaire with information on their household infrastructure; parents' educational attainment; incentives from the family to pursue an education; time watching TV, on the internet, reading books, and doing homework; if they already dropped out or repeated a grade; and if they did kindergarten.

Data from the Census of Education and *Prova Brasil* are used to calculate the National Education Development Index (IDEB), the most important educational indicator in Brazil, that monitors students' grade promotion and learning levels.[21] State and municipal governments use this indicator to monitor the improvement in the quality of public education

[19] Schools with at least 20 students enrolled in fifth or ninth grade. Proficiency tests are also applied to students in the last grade of high school.

[20] SAEB stands for *Sistema de Avaliação da Educação Básica.*

[21] IDEB stands for *Índice de Desenvolvimento da Educação Básica.* To compute the index, the students' Portuguese and math performance, on the SAEB scale, is transformed into a proficiency ranging from 0 to 10. The index is then multiplied by the students' grade promotion rate (on a scale from 0 to 1) to obtain the IDEB at school, municipal, state, and country levels. For example, for primary education, the index is the product of the standardized performance of fifth-graders and the average grade promotion from first to fifth grade.

in Brazil and to compare the schools' performance within and between municipalities.[22]

Since *Prova Brasil* is applied every two years, I use data from 2009, the year the schools were closed, and 2007, the pre-intervention year. The focus of the analysis is the fifth grade as this is the primary education grade for which the proficiency exam is applied.[23] The dependent variables of the study are then the fifth-grade proficiency in Portuguese and math, on the SAEB scale, and the independent variables are students', teachers', principals', and schools' characteristics from *Prova Brasil* and Census of Education. The regressions are run at the school level and I compare distinct cohorts of fifth-grade students.

1.5 Empirical Strategy

To identify the treatment effects of the 2009 school closures due to the H1N1 outbreak on fifth-graders proficiency in Portuguese and math, I explore the policy variation between and within the municipalities of the state of São Paulo.

In the year of the shutdowns, on the one hand, 206 of the 645 municipalities in the state of São Paulo had state-managed schools offering primary education, and the state government enforced the extension of children's winter break across all these schools. On the other hand, 642 out of the 645 municipalities had locally-managed schools offering primary education.[24] In thirteen of these 642 municipalities, the respective local authority adopted the same measure for their locally-managed schools. These state and locally-managed schools are the *treatment group*. In the locally-managed schools of the remaining 629 municipalities, the local authorities did not extend the winter break, providing a group from which the closed schools can be compared, the *comparison group*.

First, to estimate the impact of the shutdowns on the locally-managed schools, I explore the policy variation between municipalities under a difference-in-differences design. I then compare locally-managed schools of the municipalities whose local authorities extended children's winter break with locally-managed schools of the municipalities whose local authorities followed the

[22] FUNDEB stands for *Fundo Nacional de Desenvolvimento da Educação Básica*.

[23] I did not estimate the effects of the school closures on ninth-graders because only five of the thirteen municipalities whose local authorities extended children's winter break had proficiency scores available for the ninth grade in 2007. Therefore, the number of clusters would be too small.

[24] Census of Education, 2009.

school calendar as previously planned. For locally-managed schools, students' proficiency in *Prova Brasil* is available since 2005.[25] However, this proficiency assessment does not include all the primary education schools, mostly due to their size.[26] In order to test whether treatment and comparison groups have parallel trends before the school closures, I restrict the DiD to municipalities whose locally-managed schools have proficiency data available since 2005. This is the case for ten of the thirteen municipalities whose local authorities extended the winter break and for 469 of the 629 municipalities whose local authorities followed the school calendar as previously planned. Overall, the sample has 795 locally-managed schools that were closed and 2,568 that remained open (Table 1.2).

Second, to estimate the impact of the shutdowns on the state-managed schools, I explore the policy variation within municipalities. For the same municipality, I am able to compare its state and locally-managed schools where, therefore, children are affected by the same shocks at the municipality level. I restrict the sample to municipalities where there is at least one state and one locally-managed school offering primary education. This is the case of 112 municipalities where the local authority did not extend the winter break of their locally-managed schools (a group that I classify as $G = 1$), and for ten out of the thirteen municipalities where the local authority postponed the children's return to locally-managed schools ($G = 0$). In the first DiD, for schools located in $G = 1$, I compare 929 state-managed schools that were closed with 1,334 locally-managed schools that remained open. In the second DiD, for schools located in $G = 0$, I compare 868 state-managed schools with 759 locally-managed schools, all closed, in order to take into account the differences in the proficiency trajectory of state and locally-managed schools (Table 1.2). Finally, the third difference is a result of the subtraction of the first DiD and the second DiD.

[25] Proficiency data at the school level for state-managed schools are available starting in 2007.

[26] As described in Section 1.4, schools in which the classes have at least 20 students in the fifth grade are the ones that participate in *Prova Brasil*.

Table 1.2: Sample of the study, São Paulo (2005-2009)

(A) Difference-in-differences. Municipalities in São Paulo where there is at least one school managed by the local authority
Locally-managed schools only

	13 municipalities where the local authorities extended the winter break (G = 0)		*Other municipalities where the local authorities did not extend the winter break (G = 1)*	
	Census of Education	Included DiD	Census of Education	Included DiD
Number of:				
Municipalities	13	10	629	469
Schools	972	795	4,956	2,568
Students	126,562	118,804	301,743	239,778

(B) Triple difference-in-differences.
Locally and state managed schools

	13 municipalities where the local authorities extended the winter break (G = 0)				*Other municipalities where the local authorities did not extend the winter break (G = 1)*			
	Census of Education		Included triple DiD		Census of Education		Included triple DiD	
	State managed network	Locally managed network	State managed network	Locally managed network	State managed network	Locally managed network	State managed network	Locally managed network
Number of:								
Municipalities	10	13	10	10	196	629	112	112
Schools	900	972	868	759	1,244	4,956	929	1,334
Students	144,766	126,879	144,655	110,606	108,750	301,743	99,172	139,017

Notes: I only keep in the DiD sample, locally-managed schools with proficiency available since 2005 in order to test for parallel trends before the shutdowns. For the triple DiD, I only keep municipalities where there is at least one state and one locally-managed school. The rows *students* show the number of students enrolled in fifth grade. Source: Census of Education, 2009.

1.5.1 DiD on the sample of locally-managed schools

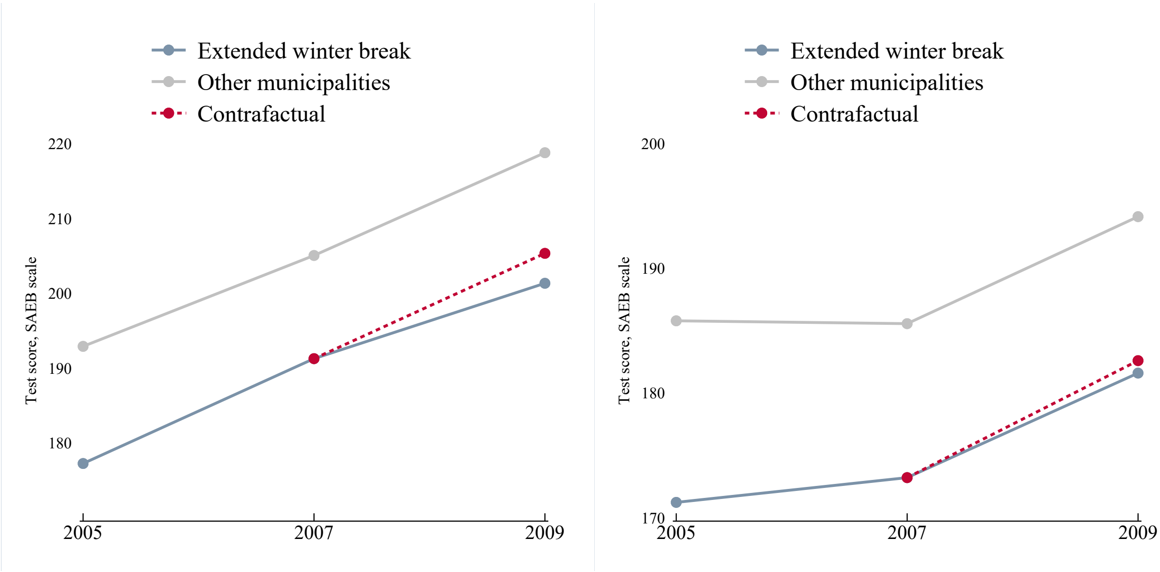
I use a difference-in-differences (DiD) approach to estimate the impacts of the school shutdowns on the locally-managed schools in São Paulo. I run the following equation:

$$y_{smt} = \alpha_0 + \alpha_1 T_{sm} + \alpha_2 H1N1 + \gamma_{DD} T_{sm} \times H1N1 + \alpha_3 X'_{smt} + \delta_m + \upsilon_{smt} \quad (1.1)$$

in which y_{smt} is the proficiency of fifth-graders in school s, in municipality m, in year t; T_{sm} is a dummy that is equal to 1 if the school s is in a municipality m that extended the winter break in 2009, and 0 otherwise; $H1N1$ is a dummy that is equal to 1 in 2009, and to 0 in 2007; X'_{smt} is a vector of control variables of students', teachers', principals', and schools' characteristics; and δ_m are municipalities fixed effects. The parameter of interest, γ_{DD}, is the average treatment effect on the treated (ATT), that is, the average effect of school closures on student learning. As schools in the same municipality may be similar, the standard error is clustered at the municipality level. As shown in Table 1.2, the number of municipalities in the treatment group (10) is significantly smaller than in the comparison group (469), therefore, the confidence interval is estimated using *wild-bootstrap* with 1,000 replications (Roodman et al. (2019)). All the regressions are weighted by the fifth-grade enrollment at the school level.

Intuitively, the effect of the winter break extension is estimated by comparing the evolution of the average test score gap between the treatment and comparison groups in 2007 (before the pandemic) and 2009 (after the pandemic). To claim the causal effects of the school closures on students' proficiency, the comparison group is expected to emulate what would have happened with the student learning in the treatment group in the absence of the school shutdowns. As shown in Figure 1.1, I find suggestive evidence that the learning outcomes in both groups were following the same trend before the H1N1 outbreak.

Figure 1.1: Students' proficiency in locally-managed schools, fifth-grade (2005-2009)



(a) Math (b) Portuguese

Note: For fifth-graders, the proficiency scale (SAEB scale) ranges from 0 to 350 in Portuguese, and from 0 to 350 in Math. The proficiency for treatment and comparison groups is the average of the proficiency scores of a sample of locally-managed schools in the state of Sao Paulo, the ones that have the performance of fifth-graders available since 2005. These schools are located in 10 municipalities of the treatment group and in 469 municipalities of the comparison group (Table 1.2). The confrafactual is calculated assuming that the average score of the treatment group would increase at the same pace as that of the comparison group in the absence of the school closures. Because the first year of *Prova Brasil* assessment was 2005, I am unable to test for parallel trends using a longer time frame. Figure A.3 shows the time-trend for grade promotion, repetition, and dropout rates. Source: Data from *Prova Brasil*/INEP.

Aiming to increase precision and account for potential time-variant confounders, I include a vector of covariates at the school level, as these variables might be correlated with both the decision of the municipality to close its schools and students' performance in the standardized tests. I use a Lasso (*Least Absolute Shrinkage and Selection Operator*) regression to select the vector of covariates, X'_{smt}, that best predict the variation in the proficiency score (Ahrens et al. (2019)).

I include the following variables in the *Lasso* regression: if the school has a science and computer lab, sports court, library, and access to the internet; instruction hours per day; the number of students per class; GDP per capita of the municipality where the school is located; and socioeconomic characteristics of fifth-graders. The vector of socioeconomic variables includes the percentage of mothers with a high school diploma; the percentage of students that have already been retained in one specific grade or dropped out of school; the percentage of white students; the percentage of female students; the percentage of students that work for pay; the percentage of students that previously studied in a private school; the percentage of students who have a computer at home; the percentage of students whose parents incentive them to study, to do the homework, to read, to not miss classes, and that talk about what happens in the school; the percentage of teachers with tenure; whether there is a lack of textbooks in the school; the quality of the textbooks; the number of enrollments in the school; and whether the students are allocated into their classes at the beginning of the school year based on their previous academic performance.[27]

Overall, the working sample for the difference-in-differences analysis consists of 795 schools with 118,804 fifth-grade students in the treatment group and 2,568 schools with 239,778 students in the comparison group (Table 1.2).

1.5.2 Triple DiD on the sample of state and locally-managed schools

In the DiD approach, one may wonder whether the influence of unobserved time-varying factors might affect differently the municipalities whose local authorities extended the winter break and the municipalities whose local authorities followed the school calendar as previously planned. If there are unobservable time-varying factors correlated with both the mayor's decision to extend the winter break and students' proficiency, the estimates based on Equation 1.1 would be biased. For example, the higher incidence of H1N1 cases in the municipalities that postponed the return to school could result in more students taking care of relatives leading to less time to study. Also, student and teacher absenteeism and psychological distress could have increased at a higher pace in municipalities more affected by the pandemic.

To overcome this identification threat, I leverage a within-municipality variation created by the school shutdown policy. To do so, I consider the municipalities with at least one state and one locally-managed school and split them into two groups. Let $G = 1$ denote the 112 municipalities in which 929 state-managed schools were closed and 1,334 locally-managed schools remained open (*comparison group*). Let $(G = 0)$ denote the ten municipalities where both 868 state and 759 locally-managed schools were closed (Table 1.2). I assume that idiosyncratic shocks at the municipal level affect students from both municipal and state schools similarly.

[27]Tables A.3, A.4, A.5, A.6, A.7, A.8, A.9 present the descriptive statistics these covariates. The model does not include variables that might have been affected by the shutdowns, such as absenteeism of students and teachers, principal managerial skills, student motivation, and teacher motivation. By including these variables, I would be more likely to underestimate the effects of the shutdowns.

Due to the differences in the learning trajectories, a simple comparison of state and locally-managed schools in $G = 1$ would likely lead to biased estimates of the impacts of the school shutdowns. On the one hand, between 2005 and 2007, locally-managed schools experienced a higher increase in average proficiency compared to state-managed schools (Figure A.4).[28] On the other hand, between 2007 and 2009, even with the shutdowns, there was a higher increase in the students' proficiency in the state-managed schools, which could suggest that the state government put forward some interventions intended to increase students' learning. In 2008, for example, the state government implemented a program aimed to increase managerial practices of schools in the bottom 5% of proficiency distribution. Despite the relatively small number of state-managed schools included in this intervention (621 out of 5,977), they are excluded from the evaluation sample to mitigate the risk of bias.[29]

To take into account that state and locally-managed schools have distinct learning trajectories, I compare state and locally-managed schools in $G = 1$ and $G = 0$ under a triple difference-in-differences design. The model accounts for three sources of variations. The first variation (Δ_1) consists of an estimate of the effects of the school shutdowns and the proficiency gap between state and locally-managed schools. It comes from the differences in learning outcomes across the state (closed), and locally-managed schools (opened) in $G = 1$, between 2007 (before the shutdowns) and 2009 (after the shutdowns). The second variation (Δ_2) consists of an estimate of the learning gap between state and locally-managed schools. It comes from the differences in learning across the state and locally-managed schools, both closed, between the same period in $G = 0$. Therefore, the third source of variation ($\Delta_1 - \Delta_2$) captures the effect of the shutdowns in the state-managed schools (Muralidharan and Prakash (2017)).

To estimate the ATT, I use the following regression equation:

$$
\begin{aligned}
y_{smt} = & \alpha_0 + \alpha_1 E_{sm} + \alpha_2 G + \alpha_3 H1N1 + \alpha_4 E_{sm} \times G + \alpha_5 E_{sm} \times H1N1 \\
& + \alpha_6 G \times H1N1 + \gamma_{DDD} E_{sm} \times G \times H1N1 + \alpha_7 X'_{smt} + \delta_m + \upsilon_{smt}
\end{aligned}
$$

in which E_{sm} is equal to 1 if school s in municipality m is state-managed and 0, otherwise. $X'_{smt}$is

[28]Columns I of Table A.11 Table A.12 and show that, even before the shutdowns, I do not have evidence that students' outcomes of state and locally-managed schools have parallel trends.

[29]In 2009, the state government had 5,977 state-managed schools offering first to ninth grade. These schools are the focus of the managerial practices intervention. Table A.15 shows the results of the triple DiD including these schools in the analysis.

a vector of control variables of students', teachers', principals', and schools' characteristics that are selected using a Lasso regression. It includes a series of variables that might be associated with how the schools reacted to the shutdowns, for example, length of the school day, managerial skills of principals according to teacher perspective, and teacher and student absenteeism. The parameter of interest, γ_{DDD}, is the average treatment effect on the treated (ATT), that is, the average effect of school closures on student learning in state-managed schools. The standard error is clustered at the municipality level. All the regressions were weighted by the fifth-grade enrollment at the school level. Table A.11 shows that with this strategy, I can control for the differences in learning trajectories between state and locally-managed schools, as the coefficient of $E \times G \times D$ is insignificant before the shutdowns (there is no significant difference in the learning gap of state and locally-managed schools in $G = 1$ and $G = 0$).[30]

To summarize:

- $G = 1$ are the 112 municipalities where the municipal governments did not extend the winter break, and where there is at least one state and one locally-managed school offering first to fifth grade. These municipalities have 929 state-managed schools with 99,172 students; and 1,334 locally-managed schools with 139,017 students.
  - locally-managed schools opened.
  - state-managed schools closed.
  - 1^{st} difference-in-differences:

$$\Delta_1 = (Y^1_{E,G=1} - Y^0_{E,G=1}) - (Y^1_{M,G=1} - Y^0_{M,G=1}) = \gamma_{DDD} + \alpha_5 \tag{1.2}$$

  - in which $Y^1_{E,G=1}$ and $Y^0_{E,G=1}$ are the proficiency of state-managed schools in $G = 1$ in $D = 1$ (2009) and $D = 0$ (2007), respectively; and $Y^1_{M,G=1}$ and $Y^0_{M,G=1}$ are the proficiency of locally-managed schools in $G = 1$ in $D = 1$ and $D = 0$, respectively.

- $G = 0$ are the 10 municipalities in which the municipal authority extended the winter break, and where there is at least one state and one locally managed school. These municipalities have 868 state-managed schools with 144,655 students; and 759 locally-managed schools with 110,606 students.

[30] Table A.17 shows the average number of state and locally-managed schools in the municipalities of the groups $G = 1$ and $G = 0$. In $G = 1$, there are on average 85 locally-managed schools and 90 state-managed schools. In $G = 0$, there are on average 20 locally-managed schools and 10 state-managed schools.

– locally-managed schools closed.

– state-managed schools closed.

– 2^{nd} difference-in-differences:

$$\Delta_2 = (Y^1_{E,G=0} - Y^0_{E,G=0}) - (Y^1_{M,G=0} - Y^0_{M,G=0}) = \alpha_5 \quad (1.3)$$

– in which $Y^1_{E,G=0}$ and $Y^0_{E,G=0}$ are the proficiency of state-managed schools in $G = 0$ in $D = 1$ (2009) and $D = 0$ (2007), respectively; and $Y^1_{M,G=0}$ and $Y^0_{M,G=0}$ are the proficiency of locally-managed schools in $G = 0$ in $D = 1$ and $D = 0$, respectively.

- Therefore, the Triple Difference-in-differences is given by: Δ_1 - Δ_2 = γ_{DDD}

The empirical strategies face a few caveats that are worth mentioning. First, since *Prova Brasil* started in 2005, I am unable to use data before 2005 to assess the plausibility of the parallel trends assumption for a more extended pre-intervention period. In this case, the triple difference design helps me deal with different time trends across municipalities by exploiting a within-municipality variation across school networks over time. Second, I do not have a panel of students. Instead, I compare distinct cohorts of fifth-graders over time and run the analysis at the school level.

1.6 Results

I find evidence that the school shutdowns during the H1N1 outbreak had a significant impact on students' learning, especially in math performance. The baseline estimates point to a decrease in Math scores equivalent to -0.21 and -0.28 of a standard deviation in locally and state-managed schools, respectively (Columns 1 and 4 of Table 1.3). The decrease in Portuguese scores is equivalent to -0.24 of a standard deviation and is restricted to the state-managed network.[31]

[31] Table A.16 shows DiD and Triple DiD estimates without adding controls for students', teachers', principals', and schools' characteristics. The DiD estimates are very similar. The triple DiD estimates are still significant and negative, however, smaller and statistically significant only at 10%.

Table 1.3: Impact of the school shutdowns on students' learning, fifth-grade (2007-2009)

Estimated decrease in Math and Portuguese Proficiency, SAEB scale																
	Math								**Portuguese**							
	DiD	DiD	DiD	Triple D	Triple D	Triple D	Triple D	Triple D	DiD	DiD	DiD	Triple D	Triple D	Triple D	Triple D	Triple D
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
H1N1	-3.26**	-3.27**	-2.75**	-4.56***	-4.47***	-4.31***	-4.26***	-4.18***	-0.76	-0.78	-0.54	-3.69***	-3.61***	-3.63***	-4.26***	-2.87***
	(1.18)	(1.22)	(1.23)	(0.85)	(0.90)	(0.91)	(0.95)	(1.28)	(0.86)	(0.91)	(0.95)	(0.73)	(0.76)	(0.75)	(0.95)	(1.05)
Wild-bootstrap p-value	0.04	0.03	0.04	0.00	0.00	0.00	0.00	0.00	0.39	0.42	0.61	0.00	0.00	0.00	0.00	0.01
95% CI	[-5.6,-0.9]	[-5.7,-0.9]	[-5.2,-0.3]	[-6.3,-2.9]	[-6.3,-2.7]	[-6.1,-2.5]	[-6.1,-2.4]	[-6.7,-1.6]	[-2.5,0.9]	[-2.6,1.0]	[-2.4,1.3]	[-5.1,-2.2]	[-5.1,-2.1]	[-5.1,-2.1]	[-6.1,-2.4]	[-4.9,-0.8]
N. schools	3912	3912	3912	5164	5164	5164	5164	5164	3912	3912	3912	5164	5164	5164	5164	5164
R2	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.6
Proficiency - Treatment Group before the school shutdowns (2007)																
Mean	193.38	193.38	193.38	195.3	195.3	195.3	195.3	195.3	175.44	175.44	175.44	177.88	177.88	177.88	195.3	177.88
Sd	15.44	15.44	15.44	16.09	16.09	16.09	16.09	16.09	15.4	15.4	15.4	15.46	15.46	15.46	16.09	15.46
ATT est (in sd)	-0.21	-0.21	-0.18	-0.28	-0.28	-0.27	-0.26	-0.26	-0.05	-0.05	-0.03	-0.24	-0.23	-0.24	-0.26	-0.19
Estimated increase in the percentage of students below the basic level of learning in Math and Portuguese, in %																
	Math								**Portuguese**							
	DiD	DiD	DiD	Triple D	Triple D	Triple D	Triple D	Triple D	DiD	DiD	DiD	Triple D	Triple D	Triple D	Triple D	Triple D
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
H1N1	2.58**	2.60**	2.19**	4.91***	4.88***	4.76***	4.75***	5.11***	1.19	1.19	0.96	3.42***	3.37***	3.34***	3.31***	3.00***
	(0.92)	(0.95)	(0.96)	(0.79)	(0.82)	(0.84)	(0.87)	(1.10)	(0.87)	(0.89)	(0.89)	(0.73)	(0.70)	(0.74)	(0.72)	(0.97)
Wild-bootstrap p-value	0.03	0.02	0.04	0.00	0.00	0.00	0.00	0.00	0.22	0.23	0.32	0.00	0.00	0.00	0.00	0.00
Percentage of students below the basic level of proficiency - Treatment Group before the school shutdowns (2007)																
Mean	57.49	57.49	57.49	56.36	56.36	56.36	56.36	56.36	70.74	70.74	70.74	69.62	69.62	69.62	69.62	69.62
Increase, in %	4.5%	4.5%	3.8%	8.7%	8.7%	8.4%	8.4%	9.1%	1.7%	1.7%	1.4%	4.9%	4.8%	4.8%	4.8%	4.3%
(A) Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(B) Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(C) % teachers in both networks	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes
(D) % teachers management	No	No	Yes	No	No	Yes	Yes	Yes	No	No	Yes	No	No	Yes	Yes	Yes
(E) School FE	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No	Yes

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the intervention of managerial practices. (E) schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented managerial practices intervention.

The magnitude of the estimates is higher than the available evidence on the impacts of school shutdowns on primary students' learning during a recent pandemic outbreak. Amid the Covid-19 pandemic, Maldonado and De Witte (2020) find that the 9-week school closure in Belgium reduced students' math and Dutch performance by 0.19 and 0.29 of a standard deviation, respectively. Also during Covid-19, Tomasik et al. (2021) show that the 9-week shutdown in Switzerland decreased math and reading performance by 0.20 of a standard deviation. Therefore, amid the Covid-19 outbreak, school shutdowns in Belgium and Switzerland were three times longer than what was experienced in Brazil during the H1N1.

The more pronounced effects among Brazilian students might be associated with the significant contrast between the students' learning levels in Brazil and OECD countries. The 2018 Programme for International Student Assessment reveals that, at the age of 15, students from Belgium and Switzerland have a reading performance almost 20% higher than students in Brazil. In math, their performance is more than 30% higher. Therefore, my higher estimates for the impact of school closures might be partially explained by the reasonable assumption that students with lower baseline levels of learning face more severe consequences. Students in Brazil have fewer resources to cope with the shutdowns, such as computers and access to the internet, especially considering that the H1N1 occurred in 2009, more than 10 years before Covid-19, when access to broadband internet and electronics was more limited and expensive. In addition to that, in 2009 there was no remote learning, whereas this measure was adopted during Covid-19. Also, children in Brazil have parents with lower educational attainment to incentive them to continue studying their textbooks while schools were closed.

The stronger negative effects on math are in agreement with other estimates of school shutdowns on learning outcomes (Kuhfeld et al. (2020), Thum and Hauser (2015), Baker (2013), and Cooper et al. (1996)). The available evidence suggests that out-of-school enrichment activities during school breaks tend not to focus on math, which is associated with both math anxiety and new instructional methods that differ from what parents themselves learned. In this context, instead of math, families tend to focus more on the promotion of literacy skills at home, by the reading of books, for example (McCombs (2011), Murnane (1975), Bryk and Raudenbush (1989), Allinder et al. (1992), Harris and Sass (2009)). Therefore, these factors help to explain why math skills depreciate faster than reading ones, why school shutdowns have larger effects on mathematics than on reading, and inform policies aimed at mitigating learning losses caused by school closures.

To get a better sense of the magnitude of the treatment effects, I convert the estimates into expected years of schooling. To do so, I use as a benchmark the ideal learning gain between the fifth and ninth grades in the Brazilian school system. An average student who does not repeat a grade should experience an annual increase in proficiency of 20 points in the national standardized exam, on a learning scale that ranges from 0 to 325 for Portuguese and 0 to 350 for math (Alves et al. (2016)). Therefore, the estimated drop in math performance of -3.3 and -4.6 points corresponds to at least six and nine weeks of learning loss in locally and state-managed schools, respectively.[32]

Given the relatively large effects of the winter break extension on learning loss, I test whether the estimated drop in test scores increased the percentage of students below the basic learning level according to the SAEB scale (Table 1.3). The baseline estimates point to an increase in the percentage of students below the basic level in math proficiency of 2.6 and 4.9 percentage points in locally and state-managed schools, respectively (a rise of 4.5% and 8.7%, compared to baseline levels). For Portuguese, the point estimate is 3.4 percentage points in state-managed schools (a rise of 4.9%, compared to baseline levels).

The relatively small increase in the proportion of students below the basic learning level suggests that the adverse impacts of the school closures were disproportionately higher among those who already lagged behind even before the pandemic. In fact, in 2007, almost 60% and 70% of the students in affected schools did not achieve basic levels of learning in math and Portuguese, respectively.[33] If the more detrimental effects are concentrated on this group of students, school shutdowns likely widened inequalities between low and high-performers.

To further investigate whether the impacts of the school shutdowns are more pronounced at the lower tail of the learning distribution, I estimate unconditional quantile treatment effects using the Athey and Imbens (2006) changes-in-changes (CiC) estimator.[34] For the state-managed

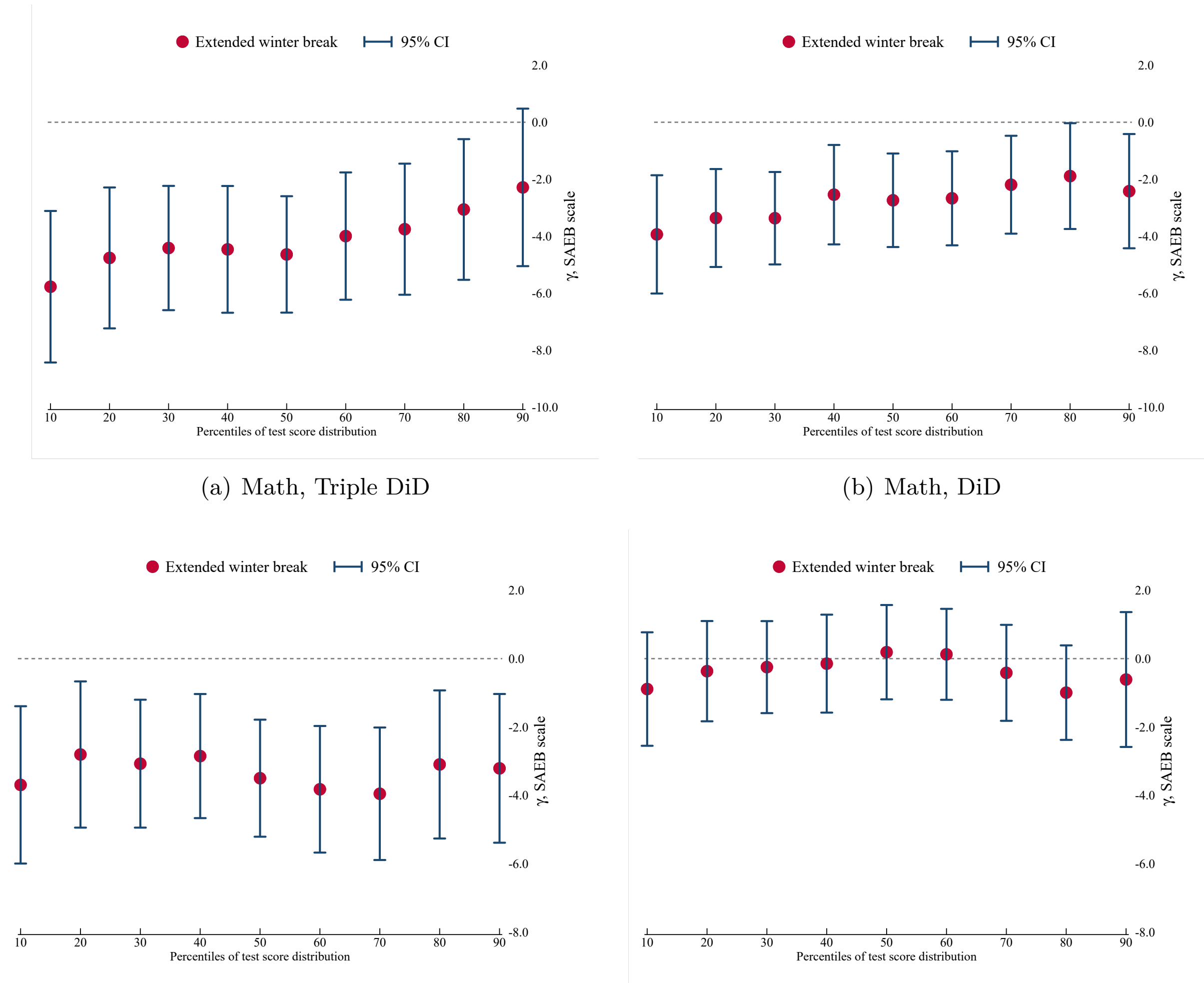
[32] 200 school days $\times$ $3.3/20 = 32$ days, or 6.6 weeks (considering five school days a week). 200 school days $\times$ $4.6/20 = 46$ days, or 9.2 weeks (considering five school days a week).

[33] Average of students with insufficient performance in 2007. The average considers all state and locally-managed schools affected by the shutdowns.

[34] I use the command *cic* in Stata to perform the estimation. This command allows the estimation of the confidence intervals with bootstrapped standard errors. However, it does not perform the *Wild Bootstrap* as defined by Roodman et al. (2019) that accounts for a small number of clusters.

networks, the impact of the shutdowns on math performance is not statistically significant in the top 10% of the distribution (Figure 1.2).

Figure 1.2: Impact of the school shutdowns by percentiles, fifth-grade (2007-2009)



(a) Math, Triple DiD

(b) Math, DiD

(c) Portuguese, Triple DiD

(d) Portuguese, DiD

Note: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. 95% Confidence Interval. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). Figures (a) and (c) show the estimates for the triple DiD. Figures (b) and (d) show the estimates for DiD. All regressions are weighted by fifth-grade enrollment at the school level. Standard errors clustered at the municipality level. The controls are municipal fixed effects, students', teachers', schools', and principals' characteristics, the percentage of teachers that work in a state and a locally-managed school at the same time, and the percentage of teachers that work in a state-managed school that implemented the managerial practices intervention. The triple difference specifications also include schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented managerial practices intervention. I perform the estimates using the Stata command *cic* (changes in changes) proposed by Athey and Imbens (2006).

The more detrimental effects in the state-managed network might indicate that locally-managed schools could better respond to the students' needs when schools reopened. Whether the impact of school closures on learning varies with the school's administration level is still an open question in the literature, there are several papers that point to the benefit of a decentralized management on the provision of education (Galiani et al. (2008), Faguet (1999), Jimenez and Sawada (1999), Filmer (2002), King and Ozler (2000)). Having a policymaker close to the population helps the identification of households' needs. In the context of school shutdowns, parents of children enrolled in locally-managed schools have lower transaction costs putting pressure on the local authority administration to take actions aimed to help students catch up with the school curriculum.

Finally, I also assess the effects of the school shutdowns on grade promotion, retention, and dropout. I find evidence that in locally-managed schools, teachers might have been more condescending to the students affected by the shutdowns. Their retention rate is estimated to be 1.3 percentage points lower compared to the comparison group, which is equivalent to a decrease of 9.6% (Table A.14).

1.6.1 Robustness

Locally-managed schools

In the municipalities that opted not to extend the winter break, 11% of the contracted teachers in their locally-managed schools also worked for the state-managed network, whose schools were closed.[35] Therefore, one may argue whether the impact of the shutdowns is underestimated as it is possible that part of these teachers opted not to lecture while state schools remained closed, making students from the locally-managed network miss school content even though the municipal government opted not to extend their winter break. To deal with this, I add as control the percentage of teachers working in a state and a locally-managed school (columns 2 of Table 1.3). The estimates are very similar to the baseline specification presented in columns 1 of Table 1.3.

As pointed out in Section 1.5, the DiD analysis is run on a sample of locally-managed schools. Therefore, none of these schools were included in the state government's managerial practices program implemented in 2008 and 2009. However, there are teachers of the locally-managed network that also worked in a state-managed school where this intervention was implemented. This is the case for 5.4% of teachers in locally-managed schools that extended the winter break,

[35] Census of Education, 2009.

and for 1.3% of teachers in the comparison municipalities. Even though the percentages are small, if these teachers took better managerial practices to the locally-managed schools, the estimates are likely to be biased. I then add as control the percentage of teachers in each one of the locally-managed schools that also work in a state-managed school that implemented the state-government intervention. The point estimate is lower in absolute terms as shown in columns 3 of Table 1.3, but not statistically different from the baseline specification (columns 1 of Table 1.3).

State-managed schools

As discussed in Section 1.5, the triple difference regression considers the sample of state and locally-managed schools in the municipalities that opted to extend the winter break ($G = 0$) and in the municipalities that followed the school calendar as previously planned ($G = 1$). To not confound the school shutdowns with the managerial practices program implemented by the São Paulo state government in 2008 and 2009, I exclude from the analysis all the 621 schools where this intervention was implemented.[36] However, 18% of the teachers of the sample of state-managed schools also worked in a participating school in the managerial practices program. Therefore, since one may argue whether these teachers could offset the impact of the shutdowns, I run an additional specification adding the percentage of teachers in this situation as a control variable. The results are very similar to the baseline estimate (columns 4 and 5 of Table 1.3).

It is interesting to notice that the first DiD of the triple differences model is run on the sample of state and locally-managed schools of the 112 municipalities ($G = 1$). The comparison here is among state schools that extended the winter break with the municipal schools that remained open (Equation 1.2 and Table 1.2). As pointed out in the previous paragraphs, there is a small percentage of teachers in these municipalities that work for both networks, which could bias the estimates. I then test a specification by adding the percentage of teachers in this situation as a control variable. The point estimate is smaller, but not statistically different than the baseline one (columns 4 and 6 of Table 1.3).

The second DiD of the triple difference model considers the schools in the 10 municipalities

[36] Table A.15 shows the results of the triple DiD including these schools in the analysis.

($G = 0$) where both state and locally-managed schools were closed (Equation 1.3 and Table 1.2). Therefore, one may wonder whether the second DiD estimate is capturing both the time trend gap between these two school networks (α_5), but also how each school network reacted to the shutdowns. On the one hand, the educational indicators of locally-managed schools are on average better than the ones of the state-managed network in $G = 1$ (Table A.4 and Table A.8). On the other hand, descriptive statistics presented in Table A.5 and Table A.9 suggest that the state-managed network is on average better than the local ones in $G = 0$. For instance, the percentage of principals who perceive teacher absenteeism as a big issue is 11 percentage points lower in state-managed schools than in locally-managed schools. On average, these state schools had longer school days than locally-managed ones (5.2 hours vs. 4.7 hours) and lower teacher absenteeism. In $G = 0$, state-managed schools also seemed to be better prepared in terms of the availability and quality of textbooks.[37]

To deal with the aforementioned issue, in addition to controlling for schools, principals, teachers, and students' characteristics, I run a model including school-fixed effects. I then control for time-invariant unobserved school characteristics, seeking to attenuate any biases accruing from heterogeneity in schools' response to the shock. In this case, the point estimates are smaller, but not significantly different from the baseline ones (columns 4 and 8 of Table 1.3.)

1.6.2 Potential mechanisms

I find evidence that a relatively small period of school shutdowns (2 to 3 weeks) caused by the 2009 H1N1 outbreak led to a significant decrease in students' proficiency in math, equivalent to at least 0.18 of a standard deviation in locally-managed schools and 0.26 in state-managed ones. For Portuguese, the estimated decrease is restricted to the state-managed network and is equivalent to at least 0.19 of a standard deviation.

The magnitude of the effect, besides reflecting the short time frame to cover the school curriculum, might also reflect other factors associated with the pandemic context. It could be that teachers, parents, and students were not entirely comfortable with in-person classes soon after the schools reopened. More stressed teachers and students could negatively affect the

[37] The percentage of teachers that classify the textbooks as great is six percentage points higher, and the percentage of principals stating there is a lack of textbooks is 26 percentage points lower.

quality of the classes, student-teacher, and peer-to-peer interactions at school. Teachers might also have been more condescending to the students after schools reopened, leading to children that would put less effort into learning. Also, there were between nine to ten weeks between the reopening of the schools and the application of *Prova Brasil*.[38] Therefore, students and teachers did not have much time to cover the curriculum before the test. Although I do not have any data to test these hypotheses directly, I highlight a few differences on the school management side and in the availability of resources that might elucidate the different responses of state and locally-managed schools.

In the following analysis, I dig into the heterogeneous effects of the shutdowns by interacting the treatment status with the ratio of students per teacher, teachers' perception of principals' managerial skills, teacher absenteeism, percentage of teachers with the adequate university degree to teach Portuguese and math, and whether the teachers always correct students' homework. I then run a regression of learning on the school closure dummy, the above-mentioned interactions, and controls aiming to assess whether the effects of the shutdowns were either mitigated or exacerbated by those factors.

I also present descriptive statistics highlighting the differences between municipalities, as well as differences between state and locally-managed schools within municipalities. By doing this, I aim to shed light on how school staff and students dealt with the negative shock.

School principals

The index of principal managerial skills ranges from 0 to 1. It is calculated based on teachers' answers of how frequently they believe that the principal pays attention to students' learning, administrative norms, and school maintenance; motivates the teachers and encourages new ideas, and takes into consideration teachers' inputs; and whether teachers trust the principal and can participate of the decisions related to their work. I test if schools in which principals have better managerial skills were better able to cope with the shutdowns.

For the state-managed network, I find suggestive evidence that better-prepared principals can attenuate the negative impacts of the shutdowns (Table 1.4). A 10% increase in principals'

[38] There are between 44 and 53 business days between the reopening of the schools (August 17, 2009) and the proficiency test (which took place between October 19 and October 31, 2009).

managerial skills is associated with a negative impact on math proficiency 0.45 points smaller on a SAEB scale, equivalent to five school days, or an effect 10% smaller than the average effects of the baseline specification.[39] For Portuguese, the same increase in principals' managerial skills is associated with an effect being 8% smaller.[40]

[39] The baseline triple difference in differences specification estimates a decrease in math performance of -4.6 points in SAEB scale (Table 1.3). As mentioned earlier, the index of principal managerial skills ranges from 0 to 1. I propose an exercise comparing a school in which the principal managerial skills is 0.91 and another in which the index is 1, therefore, 10% higher. *Ceteris paribus*, the effect of the shutdowns in a school where the principals' managerial skills reach 0.91 is equal to -3.3 points on a SAEB scale ($-8 + 5.15 \times 0.91$), whereas, in a school in which the index reaches one, the effect is equal to -2.85 ($-8 + 5.15 \times 1$). Therefore $-2.85 - (-3.3) = 0.45$. Since in 200 school days, students are supposed to increase their proficiency by 20 points, the 0.45 estimate is equivalent to 4.5 school days (200 school days $\times$ $0.45/20 = 4.5$ days).

[40] I perform the same exercise. $-5.2 + 3.3 \times 1 - 5.2 + 3.3 \times 0.91 = 0.3$. These coefficients are shown in Table 1.4. Since the baseline estimate for the decrease in Portuguese proficiency is 3.7 on a SAEB scale, the 0.3 difference is equivalent to 8%.

Table 1.4: Principals heterogeneity on the impact of the school shutdowns on students' learning, fifth-grade (2007-2009)

Effects of school shutdowns on the locally-managed network						
		Math			Portuguese	
	DiD	DiD	DiD	DiD	DiD	DiD
	(1)	(2)	(3)	(1)	(2)	(3)
H1N1	-2.75**	-3.62*	-2.69*	-0.54	-0.34	-0.13
	(1.23)	(2.04)	(1.62)	(0.95)	(1.77)	(1.38)
Wild-boostrap p-value	0.04	0.22	0.08	0.61	0.86	0.93
H1N1 versus		1.56			0.01	
principal managerial skills		(2.15)			(2.30)	
Wild-boostrap p-value		0.49			1.00	
H1N1 versus			0.1			-0.33
program to reduce dropout			(0.66)			(0.58)
Wild-boostrap p-value			0.90			0.69
N. schools	3912	3838	3857	3912	3838	3857
R2	0.8	0.8	0.8	0.8	0.8	0.8
Effects of school shutdowns on the state-managed network						
		Math			Portuguese	
	Triple DiD	Triple DiD	Triple DiD	Triple DiD	Triple DiD	Triple DiD
	(1)	(2)	(3)	(1)	(2)	(3)
H1N1	-4.18***	-7.97***	-3.72***	-2.87***	-5.21***	-1.79
	(1.28)	(1.83)	(1.34)	(1.05)	(1.65)	(1.11)
H1N1 versus		5.15**			3.31*	
principal managerial skills		(2.25)			(1.81)	
H1N1 versus			-0.28			-0.89*
program to reduce dropout			(0.53)			(0.47)
N. schools	5164	5068	5076	5164	5068	5076
R2	0.7	0.7	0.7	0.6	0.6	0.6

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the intervention of managerial practices. (E) schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented managerial practices intervention.

Teachers

Teacher absenteeism is an important issue among schools impacted by the shutdowns.[41] School principals of locally-managed schools of the municipalities that extend the winter break are almost 20 percentage points more concerned with teacher absenteeism compared to the comparison group (Table A.7). In the 112 municipalities where the state network was closed and the local one remained open, the percentage of principals that saw teacher absenteeism as a big concern was 11 percentage points higher in the state-managed schools than in the local ones (Table A.8). Therefore, one may wonder whether the effects of the shutdowns were augmented by the higher absenteeism.

For the state-managed network, I find evidence that teacher absenteeism exacerbated the impact of the shutdowns on math and Portuguese proficiency by nearly 2 points on the SAEB scale (Table 1.5). The estimated learning loss in math and Portuguese in a school in which the principal sees teacher absenteeism as a big issue is 45% and 60% higher compared to schools in which absenteeism is not a concern, respectively.[42]

Even with the shutdowns, the percentage of teachers in the state-managed network that covered at least 80% of the school curriculum is not statistically different from the locally-managed network that remained open (Table A.8). Even with the shutdowns and higher teacher absenteeism, these schools seemed to have rushed to cover the school curriculum. Also, part of the textbook was probably given by substitute teachers, to whom students were not used, making it challenging for them to keep pace with it.

[41] Teacher absenteeism is a dummy equal to 1 if the principal sees teacher absenteeism as a big issue and 0 if the principal sees it as a moderate or small issue.

[42] The estimated decrease in math and Portuguese performance in a school where teacher absenteeism is not a concern is -3.8 and -2.6, respectively. If teacher absenteeism is a concern, the estimates are equal to -5.5 (-3.8 - 1.7) and -4.2 (-2.4 - 1.8), therefore, 45% and 60% higher, respectively.

Table 1.5: Teachers' heterogeneity on the impact of the school shutdowns on students' learning, fifth-grade (2007-2009)

Effects of school shutdowns on the locally-managed network								
	Math				Portuguese			
	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
H1N1	-2.75**	-3.04	-2.28*	-6.24**	-0.54	0.12	-0.24	2.74
	(1.23)	(2.27)	(1.36)	(2.59)	(0.95)	(2.11)	(1.14)	(2.65)
Wild-boostrap p-value	0.04	0.19	0.10	0.09	0.61	0.96	0.87	0.36
H1N1 versus		0.01				-0.03		
students per teacher		(0.11)				(0.11)		
Wild-boostrap p-value		0.93				0.73		
H1N1 versus			-1.94*				-1.04	
teacher absenteeism			(1.17)				(1.27)	
Wild-boostrap p-value			0.62				0.85	
H1N1 versus teachers that				0.04				
correct Math homework				(0.03)				
Wild-boostrap p-value				0.25				
H1N1 versus teachers that								-0.04
correct Portuguese homework								(0.03)
Wild-boostrap p-value								0.24
N. schools	3912	3912	3886	3912	3912	3912	3886	3912
R2	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Effects of school shutdowns on the state-managed network								
	Math				Portuguese			
	Triple DiD	Triple DiD	Triple DiD	Triple DiD	Triple DiD	Triple DiD	Triple DiD	Triple DiD
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
H1N1	-4.18***	-3.71**	-3.81***	-15.01***	-2.87***	-2.59*	-2.43**	-6.25*
	(1.28)	(1.61)	(1.26)	(4.11)	(1.05)	(1.35)	(1.05)	(3.47)
H1N1 versus		-0.03				-0.02		
students per teacher		(0.06)				(0.05)		
H1N1 versus			-1.69*				-1.82**	
teacher absenteeism			(1.00)				(0.89)	
H1N1 versus teachers that				0.13**				
correct Math homework				(0.05)				
H1N1 versus teachers that								0.04
correct Portuguese homework								(0.04)
N. schools	5164	5164	5131	5164	5164	5164	5131	5164
R2	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the intervention of managerial practices. (E) schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented managerial practices intervention.

Length of the school day and quality of the textbooks

The fact that the state-managed network was able to cover as much school content as the locally-managed network that remained open can be also explained by the longer length of their school day (5.3 hours × 4.9 hours). The 24-minute daily difference might have helped to compensate for part of the period of the shutdowns. This difference is equivalent to an addition of three to four school days in the whole school year compared to locally-managed schools (Table A.4).[43]

However, the additional time in the classroom due to the longer duration of the classes in the state-managed schools is still inferior to the lost time in the classroom due to the period of the shutdowns. Also, I do not find any indication that the state government extended the length of the school year to compensate for the days the schools were closed. In fact, the length of the school year was five days longer in the locally-managed network (Table A.4). These statistics also suggest that teachers of the state-managed network had to rush to cover the school curriculum, making it challenging for the students to keep pace with it.

A glance at the descriptive statistics presented in Table A.3 also highlights the challenges faced by the locally-managed schools affected by the shutdowns. One of the issues that stand out is that these schools were not able to cover the school curriculum in a shorter time frame. While 55% of locally-managed schools in municipalities that did not extend the winter break covered more than 80% of the school curriculum, only 45% did so in municipalities affected by the school closures.

Data from *Prova Brasil* questionnaire show that teachers faced many barriers to covering the entire school curriculum successfully. For instance, students in treated schools had fewer hours of classes per day (4.8 hours × 5 hours). The 12-minute daily difference resulted in children from the affected schools having two days less of school content between the reopening of the schools and the proficiency test.[44] Also, I do not find any indication of an extension of school

[43] There are between 44 and 53 business days between the reopening of the schools (August 17, 2009) and the proficiency test (which took place between October 19 and October 31, 2009). Therefore, the 24-minute difference would result in 1,056 to 1,272 additional minutes of class, equivalent to 17.6 to 21.2 hours. If the average number of class hours in the state-managed network in $G = 1$ is 5.3, the estimate in terms of school days is between three and four.

[44] There are between 44 and 53 business days between the reopening of the schools (August 17, 2009) and the proficiency test (which took place between October 19 and October 31, 2009). Therefore, the 12-minute difference would result in 528 to 636 fewer minutes of class, equivalent to 8.8 to 10.6 hours. If the average number of class hours in the affected network is 4.8, the estimate in terms of school days is roughly two.

days later in the year to account for the missing days during the winter break extension. The length of the school year was actually shorter in the affected network (Table A.3).

The availability and the quality of textbooks also seem a concerning point when the locally-managed network affected by the shutdowns is compared to the unaffected one. the percentage of teachers satisfied with the textbooks was 3.5 percentage points lower in treated schools, whereas the percentage of principals complaining about the lack of textbooks was 22.5 percentage points higher. The elevated teacher absenteeism (18 percentage points higher compared to the unaffected schools) and the lack of proper high-quality textbook materials suggest that the treated schools dealt with more challenges to mitigate the learning loss caused by the school closure. Also, the higher student absenteeism (4 percentage points) in affected schools means that one missing day in school implies more loss of instruction (Table A.7).

1.7 Discussion and Conclusion

The decision to close schools during a pandemic faces a clear trade-off: on the one hand, this policy seems an efficient measure to reduce infection rates among students (Adda (2016)). On the other hand, its long-term consequences can be daunting to a whole generation of youth learners, particularly the most vulnerable.

In the state of São Paulo, Brazil, amid the H1N1 outbreak in 2009, more than half of the public primary, lower, and upper secondary schools were closed for two to three weeks, affecting more than 5.5 million students. I leverage this natural experiment to estimate the impact of school closures on the proficiency of fifth-graders in Portuguese and math. In each one of the 645 municipalities of the state, the state government and the respective municipal government share the responsibilities for the provision of public education. The way the policy was implemented, allows me to investigate the impacts of the school shutdowns on the schools managed by the state government, state-managed, and on the schools managed by the local authorities, the locally-managed.

I find evidence that the school shutdowns led to a reduction in math scores of at least 0.18 of a standard deviation in locally-managed schools and 0.26 in state-managed ones, equivalent to more than six weeks of learning loss. For Portuguese, the effects are restricted to schools managed by the state authority and reach at least 0.19 of a standard deviation, suggesting that the locally-managed schools were better able to respond to the students' needs when schools reopened. Parents of children enrolled in locally-managed schools have a lower cost of putting pressure on the local authority administration for students to catch up with the missing curriculum. The more detrimental effects in math might be associated with the fact that during school breaks, families tend to focus more on the promotion of literacy skills at home, which indicates why the quality of the schools potentially has more significant effects on math than on reading.

I compare students that were away from school for two weeks, the length of students' winter break, with students that spent between four to five weeks without going to school, which was the length of the winter break together with the school shutdowns. Since the longer the period away from school, the more detrimental the loss of learning skills, my estimates also reflect the learning slide that occurred during school breaks.

The quantile estimates on math test scores indicate that the effects of school closures were higher in the schools at the bottom of the math test score distribution. The effects being higher among low-performers help explain the relatively small increase in the percentage of students below the minimum level of math proficiency (4.5% and 8.7% in locally and state-managed schools, respectively). Hence, the results indicate that the shutdowns hit harder on those lagging behind even before the pandemic struck. The data show that these low-performer schools had higher repetition and dropout rates, suggesting that the effects were higher among vulnerable students.

A pandemic outbreak might affect children's skills in various ways. A pandemic context that involves death, income loss, unemployment, insecurity, social isolation, and increased exposure to domestic violence can have meaningful negative consequences on children's socio-emotional skills and well-being. The Global Education Monitoring Report (2019) points to the disruptive effects on learning as physical; emotional, with anxiety, fear, sadness, and lack of emotional control; and cognitive, expressed via difficulty paying attention, inability to process information, and memory problems. As emotionally nurturing environments produce more capable learners, these factors will have more profound consequences for children's cognitive development (Cunha and Heckman (2007)). Given that all these factors affect learning, the estimates could be also capturing these indirect effects caused by school closure on student learning.

For instance, the risk of contagion and the average number of hospitalizations were higher among the municipalities whose local authority imposed school closures. It could be that parents and guardians were more reluctant to send children back to school, causing them to miss more school days. After the Influenza outbreak in 1918, around 200 thousand students were absent even after the schools reopened in New York City (Meyers and Thomasson (2017)). After the Ebola outbreak in 2014, 9 out of 10 people interviewed in Ghana by the United Nations confirmed the reduction in school attendance (United Nations Development Group (2015)).[45] The fear of contagion could also have affected the leisure activities of students, potentially triggering psychological factors such as anxiety and depression with clear implications to cognitive development.

The expressive learning losses detected during a relatively small period of school shutdowns amid the H1N1 outbreak illustrate the challenges policymakers face to design public policies capable of counterbalancing the short and longer-term consequences of school shutdowns. To give context to the challenges facing the school system, I use a meta-analysis conducted by McEwan (2015) that summarizes the impacts of more than 70 randomized controlled trials of educational interventions in developing countries. According to the author, school interventions in primary education have, on average, positive effects that range from 0.05 to 0.15 of a standard deviation. The effect of the most effective intervention is smaller than my lowest estimate.

[45] The United Nations carried out an anthropological survey to capture people's practical experiences of the socioeconomic impact of the Ebola pandemic in Guinea. See the report in this link: https://ebolaresponse.un.org/socio-economic-impact-ebola-west-africa-pdf.

2 *Teaching at the Right Level*: evidence from the Program *Acelera* in Brazil

2.1 Introduction

In primary education, children build the foundation skills on which their future learning will be based (Crouch and Gove (2011)). However, more than half of the children living in developing countries are not able to read and understand a basic text by age 10.[1] In this context, some countries have implemented an intervention called *Teaching at the Right Level* targeting students that are well behind the expected level of achievement for the grade they are enrolled (Banerjee et al. (2016)). This program aims to build basic reading and math skills by grouping children, not according to their age, but according to their proficiency (Banerjee et al. (2017)). In this Chapter, I assess the effects of the program *Acelera* in the municipality of Recife/Brazil. *Acelera* is a type of *Teacher at the Right Level* that focus on students that are older than the adequate grade for their grade of enrollment and that lag behind their peers. The program aims to increase learning levels and grade promotion, and decrease dropout and age-grade distortion of primary education students.

A good set of reading and numeracy skills helps children to achieve their full potential which can lead to major life-long consequences, such as better job opportunities and breaking the intergenerational cycle of poverty (Akresh et al. (2018)). Developing countries have made huge progress in universalizing access to primary education. According to UNICEF, the net primary enrollment rate for children aged between 6 and 11 years old reaches 90%, and the percentage of them that complete this level of education is 80%.[2] Nonetheless, regarding the learning levels, there are challenges to be overcome, especially after the school shutdowns that took place during the Covid-19 outbreak. According to Donnelly and Patrinos (2021) the average learning loss amid the pandemic is equivalent to 0.17 of a standard deviation, or half of a school year.[3]

Children's low proficiency can be explained by several factors such as having a disadvantaged socioeconomic background, the lack of qualified teachers and textbooks, bad school

[1] https://www.worldbank.org/en/topic/education/brief/what-is-learning-poverty.
[2] https://data.unicef.org/topic/education/primary-education.
[3] The authors analyze thirty-six studies based on data from twenty countries.

infrastructure, and a high opportunity cost of going to low-quality public schools versus joining the labor market and complementing their households' income. Also, overall, children have to learn a demanding curriculum, regardless of their level of preparation. Since knowledge is cumulative, those who fail to achieve the right set of skills in elementary education will struggle to learn in the later grades and may never catch up (Crouch and Gove (2011)). This scenario might lead to grade retention or dropout, and, consequently, an increase in age-grade distortion, characterized by children being older than the adequate age for their grade of enrollment.

Teaching at the Right Level is an educational intervention that groups children lagging behind their peers into a new class, either for a few hours during the school day or after their regular classes. The main benefit of the intervention is to provide a class where the curriculum is tailored to match the learning levels of the students who cannot follow the standard curriculum. Children might feel more comfortable in classes where their peers have similar performance, and where they are able to keep pace with the subjects being taught (Banerjee et al. (2007)). The intervention can even benefit children who are not directly targeted. Removing low-performers drops the pupil-teacher ratio in the students' original class, and may allow teachers to focus on more advanced material.

In Brazil, *Acelera* is an intervention developed by the Ayrton Senna Institute, a non-profit organization. The program was first implemented in 1997, the year in which more than 40% of primary education students in the country were at least one year older than the adequate age for their grade, dropout rates reached more than 9% and grade retention was 13%.[4] Since then, *Acelera* has reached almost 25% of the Brazilian municipalities, one million children, and 33 thousand educational professionals.[5] The intervention provides training to the teachers so they can introduce new textbooks tailored to meet students' needs. The program differs from the conventional *Teaching at the Right Level* initiatives as it consists in allocating program beneficiaries into a new class for the whole school year, instead of just for a few hours.

The municipality of Recife, in partnership with the Ayrton Senna Institute, has been implementing *Acelera* in primary education schools under the management of the municipal

[4]Census of Education, 1997.

[5]Check the numbers in the link: https://institutoayrtonsenna.org.br/pt-br/como-atuamos/acelera.html.

authority since 2010 (locally-managed schools).[6] Before the beginning of the school year, the Municipal Department of Education selects schools to implement the intervention based on the number of potential beneficiaries and the availability of additional classrooms to accommodate them. The number of potential participants inside each school is based on the number of children that are at least one year old than the adequate grade for their grade of enrollment and on students' learning levels. Once the schools are selected, teachers and principals are in charge of selecting the participant students based both on their age-grade distortion and on the results of a proficiency assessment applied at the beginning of the school year. The intervention is then introduced to the student's parents, who can decide whether their children will participate or not.

By 2018, more than eight thousand students have participated in *Acelera*. 85.4% of the beneficiaries are either enrolled in fourth or fifth grade and around two-thirds of the locally-managed schools in Recife have implemented the program in at least one year between 2010 and 2018 (Table 2.2). Among the participating students, 86.3% are at least one year older than the adequate age for their grade. Considering all the students with at least one year of age-grade distortion that are enrolled in locally-managed schools in Recife, nearly 8% of them are included in *Acelera*, meaning that there is a significant number of non-participant students that could be potential beneficiaries, but end up not participating.[7] Since I do not have access to the results of the proficiency assessment applied at the beginning of the school year to identify potential program participants, I define the eligible group as the students that have at least one year of age-grade distortion (as 86.3% of program participants attend this criterion).[8] Also, not all the students with age-grade distortion in schools offering *Acelera* are included in the intervention, which is due to the limited capacity of the program (Table 2.2).

Acelera aims to increase students' performance in reading and math and grade promotion, and decrease dropout and age-grade distortion. According to Ayrton Senna Institute, the ultimate

[6]The schools in the municipality of Recife under the management of the state authority do not implement *Acelera*. The state government of Pernambuco, where the municipality of Recife is located, does not have a partnership with the Ayrton Senna Institute to implement *Acelera* in these schools.

[7]Between 2010 and 2018, the locally-managed schools of Recife registered 92,228 primary education enrollments of students with at least one year of age-grade distortion (57,961 + 26,998 + 7,269). The 7,269 *Acelera* enrollments with at least one year of age-grade distortion represent nearly 8% of 92,228 (Table 2.2).

[8]The municipal government of Recife only shares the EMPREL and SAEPE datasets, which do not contain data on the tests applied at the beginning of each school year to identify potential *Acelera* participants. Each school applies a distinct test and they do not set up an organized dataset that would allow me to check students' performance.

goal of the intervention is to make children jump up to two grades, so third-graders could jump up to fifth grade, fourth-graders to sixth grade, and fifth-graders to seventh grade, boosting the reduction of the age-grade distortion. In this Chapter, I test the hypothesis of whether *Acelera* achieves its goals by grouping children according to their learning levels, providing textbooks that are tailored for pupils that lag behind their peers, and establishing a new daily teacher routine.

To perform the analysis, I set up a rich dataset at the student level by combining three sources of data from 2010 to 2018. First, the EMPREL dataset contains information at the student level, such as grade of enrollment, date of birth, type of enrollment (*Acelera* or regular education), and status of the student at the end of the school year (whether the student is promoted to the next grade, retained, or dropped out).[9] Second, the SAEPE dataset also contains data at the student level, such as their performance in reading and math in the standardized test applied by the state government of Pernambuco, of which the capital is the municipality of Recife. Between 2010 to 2015, the primary education students included in the assessment were third and fifth graders. Starting in 2016, second and fifth graders have to answer the test. Third, the Census of Education from which I use the dataset at the school level to check schools' infrastructure, such as a library, computer and science labs, and provision of meals. I perform one extensive work of data cleaning to deal with inconsistencies such as (i) duplicated observations, (ii) correction of students' status at the end of the year based on their grades of enrollment at the years t and $t-1$, (iii) typos in students' and mothers' names, (iv) absence of the date of birth in a specific year or distinct dates of birth for the same student depending on the year, (v) distinct codes of enrollment for the same student depending on the year, among others.[10] The dataset has more

[9] Regular Education is the term used to describe the educational experience of typically developing children, the ones that follow the standard curriculum and that do not require special treatment.

[10] (i) Students that are enrolled in two distinct schools in the same year, that are enrolled in two distinct grades in the same school and year, that have two enrollments in the same school and year, but with distinct status at the end of the year, etc. (ii) There are cases in which the student is enrolled in the third grade in t and in fourth grade in $t+1$ but the status at the end of the year t is dropped out or retained, whereas it should be promoted to the next grade; or students that are enrolled in third grade in t and $t+1$ for which the status at the end of the year t is grade promoted, whereas it should be retained or dropped out. (iii) Students with typos such *SILVA* as the last name in t and *SILV* in $t+1$. Correcting students' names is an essential task as I use this variable to merge EMPREL and SAEPE datasets. Also, correcting mothers' names is crucial as I use this variable, together with the student's name and date of birth, to guarantee that the same student keeps the same enrollment code during the whole period of analysis. (iv) Students whose date of birth is listed as 02/03/1998 in t and as 03/02/1998 in $t+1$, or that have a date of birth in all the years with the exception of one. Correcting dates of birth is also very important as I use this variable to calculate the difference between students' age and the adequate age of their grade of enrollment. (v) Especially for students who migrate to a new school, some are then listed with a new enrollment code. Following the student over the years is crucial for the analysis as I want to assess whether the intervention affects their educational outcomes over time. My code is available on GitHub at the following link: https://github.com/vivianamorim/seliga-acelera-recife.

than seven hundred thousand enrollments between 2010 and 2018.

I assess the effects of *Acelera* on five educational outcomes: performance in reading and in math, grade promotion, age-grade distortion, and dropout. I compare program participants, the *treatment group*, with students that, although eligible to participate, are not included in the intervention, the *comparison group*. My identification strategy consists of a difference-in-differences approach, in which I include several controls at the student and at the school level, as well as school-fixed effects to disentangle the impacts of the intervention from other factors associated with both educational outcomes and program participation. In addition to that, to have more comparable groups, I run a propensity score matching on the sample of eligible students and only keep in the sample students in the common support, that is, students with similar probability of treatment based on their observable characteristics. I show the results for the whole sample of locally-managed schools, that is, schools that offer *Acelera* (where I have treatment and comparison groups) and schools that offer only Regular Education (where I only have the comparison group). I also show the results considering only schools that offer *Acelera* (where I compare treatment students with students that, although enrolled in the same school and eligible to *Acelera*, are not included in the intervention).

To estimate the impact of *Acelera* on grade promotion, dropout, and age-grade distortion, the data available allows me to follow a panel of first to fifth-graders. I find evidence that the program not only increases grade promotion in the year of the treatment but also when students return to regular education. The results show an increase in grade promotion of 16 percentage points, equivalent to a jump of 22%. As a consequence, the age-grade distortion decreases by 10 percentage points, equivalent o a decrease of 17%. No effects are found on dropout rates.

Due to the availability of the data, to estimate the effects of *Acelera* on students' proficiency in reading and math, I rely on students that join the intervention in the fourth grade. From 2010 to 2015, third and fifth graders are the students included in the standardized proficiency exam (SAEPE) and, therefore, the ones that have proficiency scores in reading and math on the standardized test applied by the state government of Pernambuco. Hence, by focusing the analysis on fourth graders, I can compare the proficiency of participants and non-participants before the intervention, when they are in third grade, and after the intervention, when in fifth grade. In this sense, I restrict the sample of analysis to fourth graders whose proficiency data is available for both third and fifth grades. I normalize the students' proficiency so that the treatment effect could be measured in terms of standard deviations (SD).

I do not find evidence that the program increases students' proficiency in reading and math. However, this result should be interpreted with caution. 42% of fourth graders in *Acelera* are promoted to six-grade after program participation. As a consequence, these students do not participate in the standardized proficiency assessment applied to fifth grade. *Acelera* students that are able to jump to the sixth grade are not in the analysis as the treatment group has fourth graders that have proficiency scores for both the third and fifth grades. Under the reasonable assumption that the ones that jump two grades are among the best performers, I end up underestimating the true impact of the intervention.

In this Chapter, I provide treatment effects estimates for the biggest *Teaching at the Right Level* intervention ever implemented in Brazil. I perform one extensive work of data cleaning that is publicly available for reproducibility and further extensions to the community of researchers. To my knowledge, this is the first non-experimental analysis that rigorously assesses the impact of *Acelera* on students' proficiency in reading and math using a rich dataset at the student level. Oliveira et al. (2019) employs data at the school level for this task, however, since only 24.6% of eligible students in schools offering *Acelera* are included in the intervention, the dataset at the student level is the most adequate one to estimate the impact of the intervention on students' performance. Although the ultimate goal of *Acelera* is to make students jump up to two grades to boost the reduction of the age-grade distortion, students' performance in standardized exams is the most recommended educational outcome to assess the impact of an educational program. By looking only at grade promotion, one can super estimate the impact of the program. There is the possibility that teachers promote the students to the next grade solely because they are program beneficiaries or due to principals' or educational managers' demands, instead of them absorbing school content and being prepared for the next grade. Also, the available evidence on *Teaching at the Right Level* mainly uses students' performance in standardized exams as their main dependent variable. Banerjee et al. (2007), Gorard et al. (2017), Duflo et al. (2020),Fryer Jr and Howard-Noveck (2020) find evidence of impacts equivalent to 0.6, 0.24, 0.15 and 0.09 of a standard deviation, respectively. My work then contributes to a growing literature on remedial education policies.

Apart from this introduction, this chapter is organized as follows: Section 2.2 presents the related literature. Section 2.3 summarizes *Acelera*. Section 2.4 introduces the data available to perform the analysis. Section 2.5 discuss the empirical strategy. Section 2.6 presents the preliminary findings. I then conclude in Section 2.7.

2.2 Related Literature

Students with low proficiency levels face more difficulties in absorbing school content, which might lead to grade repetition, dropout, and, as a consequence, an increase in age-grade distortion (defined as the percentage of students that are at least one year older than the adequate age for the grade they are enrolled in).[11] A higher percentage of students in this situation increases classroom heterogeneity, making it challenging for teachers to deal with students of such diverse ages and weakening peer-to-peer interaction. For the government, higher levels of age-grade distortion increase the inefficiency in the use of public resources as students remain in school for longer than expected.

The lack of foundation skills is the major source of grade repetition, poor performance, and high school dropout (Somers et al. (2010)). Hernandez (2011) find evidence that almost 20% of third-graders that do not read proficiently end up not graduating from high school on time, a rate that is four times higher than the one of proficient readers.

School policies can play an important role in increasing children's cognitive skills (Hanushek and Woessmann (2012)). Therefore, countries can boost their economic growth and reduce inequalities by expanding access to primary education and improving its quality. It is well documented that students' proficiency in reading and math are strongly correlated with labor productivity (Hanushek and Kimko (2000), Ciccone and Papaioannou (2009)). Indeed, Hanushek and Woessmann (2012) estimate that an increase of one standard deviation in cognitive skills of a country's workforce is associated with a roughly two percentage points increase in annual per capita GDP.

The *Balsakhi Program* is a world-recognized *Teaching at the Right Level* intervention

[11] In primary education, the adequate age is 6 for first grade, 7 for second grade, 8 for third grade, 9 for fourth grade, and 10 for fifth grade.

implemented in India. The program hires young women from children's local communities to teach basic literacy and numeracy skills to students falling behind their peers. Children included in the intervention are taken out of their main class for two hours a day and grouped into a new class that has 15 to 20 children.[12] The intervention has a standardized curriculum that focuses on the competencies that children should have learned in their two first grades. Banerjee et al. (2007) find evidence that the *Balsakhi Program* has a high impact on the proficiency of the treated students, an estimate that ranges from 0.6 to 1 standard deviation. The authors also test whether the reduction in the pupil-teacher ratio, caused by the weakest peers being removed from the classroom, benefits students not included in the program. The authors do not find evidence that this is the case, corroborating that inputs, when not accompanied by a changing pedagogy, might not help.

The *Switch-on Reading* is a 10-week intervention designed for seven graders with low reading performance in England. Students are removed from their main class and taken to a 20-minute session tailored to improve their reading comprehension and fluency. For each student, one school staff is assigned, commonly a teacher assistant, to whom they read four books chosen according to their needs. Gorard et al. (2017) find evidence that the program leads to an increase in reading skills by 0.24 of a standard deviation, equivalent to three months of learning. The authors also investigate the effects of the intervention at different levels of proficiency distribution prior to the treatment. Their results suggest that the low performers benefit more, with an effect size of 0.39 of a standard deviation, equivalent to five months.

The *High-Dosage Reading Tutoring* targets middle school students in New York City public schools. The students are grouped by 2.5 hours after the school day for supplemental classes. For a subset of them, the intervention provides a four-on-one reading tutoring section that lasts between 45 to 60 minutes. Fryer Jr and Howard-Noveck (2020) find evidence that the intervention increases the school attendance of black students by 2 percentage points and reading scores by 0.09 of a standard deviation.

The *Enhanced Reading Opportunities* is a literacy program implemented in England that targets ninth graders whose reading skills are at least two years below their grade level. Students, in addition to their regular English classes, are allocated to a 45-minute daily class with other

[12]According to Banerjee et al. (2007), their whole school day is about 4 hours.

10 to 15 children. In those classes, teachers work on students' motivation, reading fluency, vocabulary, comprehension, phonics, and writing. Somers et al. (2010) find evidence that the program improves students' reading comprehension by 0.09 of a standard deviation.

The *Teacher Community Assistant Initiative* provides two remedial education interventions for primary students in Ghana. In each one of them, one teacher assistant works with remedial learners. Duflo et al. (2020)'s findings suggest that these programs increase students' test scores by 0.15 of a standard deviation after two years of exposure.

Overall, the available literature focuses on interventions that take students out of their main class for a few hours during their school day, or that provide extracurricular activities after their school hours. Also, most of the programs are implemented for a few months or weeks of the school year. Differently from the previous interventions, *Acelera* selects low performers to be allocated into a new class for the whole school year.

I am aware of only one quantitative study on the impact of *Acelera*. Oliveira et al. (2019) perform a difference-in-differences approach using data at the student level and find evidence that the program increases grade promotion of primary students by 15 percentage points. The authors do not assess the impact of the intervention on dropout and repetition. Using data at the school level, the authors do not find evidence that the intervention affects students' proficiency in reading and math. The main caveat of this analysis is that less than 25% of the students with age grade distortion in schools offering *Acelera* are included in the intervention. Therefore, most of the average proficiency at the school level comes from students enrolled in regular education. I enrich the analysis by exploring the effects of *Acelera* on students' proficiency using a dataset at the student level. Since there is no unique enrollment code that allows me to merge EMPREL (where I identify program beneficiaries) and SAEPE (where I have students' performance), I rely on students' names, dates of birth, grade, and school of enrollment to perform the merging of the datasets. This task is one of the reasons that highlight the importance of extensive data cleaning. Also, to accommodate eventual typos in students' names not corrected in the data cleaning, then increasing the percentage of matched students, I incorporate a technique that performs the merging based on similar text patterns. I am able to find 70% of the students included in the proficiency test in the EMPREL dataset.

Oliveira et al. (2019) compare program *Acelera* participants with all the remaining students with at least one year of age-grade distortion not included in the intervention. In addition to that, to have more comparable groups of participants and non-participants, I perform a propensity score matching based on students' grades, sex, the color of the skin, level of age-grade distortion, and whether the student repeated in the previous year. By restricting the analysis to students in the common support, I am able to compare similar students, with the difference that one is included in the intervention and the other is not.

2.3 The Intervention: *Acelera*

The program *Acelera* targets mostly low performers primary education students who are at least one year older than the adequate age for their grade. Since 15% of the program participants do not have age-grade distortion, the program also selects students' that lag behind their peers, although are not older than expected for the grade. The intervention is designed to promote learning practices that are adequate to students' levels of proficiency, making it possible to catch up with the right set of skills for the grade they are enrolled in. By doing that, the program aims to increase learning levels and grade promotion, and decrease dropout and age-grade distortion. The ultimate goal of *Acelera* is to make students jump up to two grades. In this case, third graders could jump to fifth grade, fourth graders to sixth grade, and fifth graders to seventh grade, boosting the reduction of their age-grade distortion.

Acelera is an intervention designed by the Ayrton Senna Institute (IAS), a non-profit organization founded in 1994, that provides technical support to educational policies aimed to increase the quality of education in Brazil. *Acelera* was first implemented in 1997 by 15 Brazilian municipalities. 25 years after, almost one million students and 33 thousand educational professionals participated in this intervention. Since 2010, *Acelera* has been implemented in Recife, a municipality of the state of Pernambuco in Brazil.

According to the 2020 Census of Education, the municipal government of Recife is in charge of the management of 205 primary schools (locally-managed schools).[13] Not all of them offer

[13]In the municipality of Recife, in the state of Pernambuco/Brazil, the state government, represented by the Governor, and the municipal government, represented by the mayor, provide public primary education. According to the 2020 Census of Education, the state government has two primary schools under its management in the municipality (state-managed schools). The main roles of the state and municipal Departments of Education are

Acelera, instead, at the beginning of the school year, the Municipal Department of Education selects schools to implement *Acelera* based on the number of potential beneficiaries and the availability of additional classrooms to accommodate them. The number of potential participants inside each school is based on the number of children that are at least one year old than the adequate grade for their grade of enrollment. Once the schools are selected, teachers and principals are in charge of selecting the participant students based on their age-grade distortion and on the results of a proficiency assessment applied at the beginning of the school year. The intervention is then introduced to the student's parents, who can decide whether their children will participate or not.

Acelera participants are allocated into a new class for the whole school year. The class can have students from first to fifth grade. The program is managed by a municipal coordinator, indicated by the Municipal Department of Education, that has the responsibility to send monthly reports to the Ayrton Senna Institute, monitor students' and teachers' frequency, and ensure that the minimum length of school days is met. In addition to the municipal coordinator, local supervisors are assigned to monitor four *Acelera* classes. They attend one lecture per month to help teachers with class planning and new pedagogic practices.

Teachers participate in one training provided by the Municipal Department of Education and the Ayrton Senna Institute. The training is called *Capacitar* and consists of twelve video classes of 30 minutes each. The videos contain examples of situations that can be faced by the students during Portuguese, Mathematics, Sciences, History, and Geography classes. The recommendation is that the teachers watch the videos together, and plan their classes accordingly. Teachers are encouraged to engage students in short-term and small projects, of one or two days, in order that they can actively participate in the learning process and work on their self-esteem as they conclude their activities.

Every two weeks, teachers have meetings with their local supervisors to discuss program implementation, share experiences, and plan their classes based on the students' needs. The Ayrton Senna Institute provides textbooks to the students, the pedagogic material to develop

---

to implement programs aimed to improve the students' learning or to reduce dropout, grade retention, and age-grade distortion, as well as hire teachers, provide textbooks, appoint principals, and finance school infrastructure. This chapter does not contain information on schools in the municipality of Recife that are managed by the state government of Pernambuco, since the state authority does not implement *Acelera* in schools under its management.

the activities specified in the books, dictionaries, maps, magazines, and technical manuals for teachers and local supervisors. Table 2.1 introduces the theory of change of *Acelera* .

Acelera faces four main challenges. The first one is the heterogeneity of the classes which can have students enrolled in distinct grades and with diverse levels of age-grade distortion. The second is the stigma of participating in the intervention as students are taken out of regular education for the whole school year and could be identified as the lower performers. The third is that participating students can be the more challenging ones and one may wonder whether the best teachers would be willing to teach them. The fourth is the high levels of teacher absenteeism and turnover.

It is important to point out that, in 2001, the Ayrton Senna Institute created another program called *Se Liga*, focusing on students that do not know how to read and write. The IAS designed this intervention after observing that part of the students that are assigned to *Acelera* could not read or do basic math. *Se Liga* aims to literate students and prepare them to participate in *Acelera* or return to regular education. Besides *Acelera*, the municipality of Recife has been implementing the *Se Liga* since 2010. In this work, I do not study the effects of this intervention.

Table 2.1: Theory of Change of *Acelera*

Inputs	Activities	Outputs	Outcomes	Final outcomes
Physical space in the schools to set up *Se Liga/Acelera* classes.	Teachers map potential eligible students based on age distortion and proficiency.	Acelera: teachers cover 4 textbooks with 30 classes each. The books have Portuguese, Math, Geography, History, and Science content.	Classes with fewer students.	Increase of approval.
Teachers to be in charge of these students.	The Secretary of Education selects the schools to offer the intervention.	Se Liga: teachers cover the 42 classes of the student's textbook. This book has Portuguese and Math content.	Classes more homogeneous in terms of proficiency.	Decrease of dropouts.
Mediators to monitor the program's implementation. One mediator is in charge of four classes.	The Secretary of Education and the school's pedagogic team selects the students who will be invited to participate.	Monthly delivery of three and four extracurricular books for Se Liga and Acelera students, respectively.	School curricula in accordance with student's needs and their level of learning.	Decrease in age distortion.
Technical staff in the Secretary of Education to train teachers and supervise the program.	The interventions are introduced to parents, and students are invited to participate.	A school year with a minimum of 200 days and 800 hours.	Increase in student's self-esteem.	Increase of proficiency.
Online teacher training provided by IAS.	Teacher's training before the beginning of the school year.	Effective use of the time. The classes are organized as follows: welcoming of students, reading, correction of the homework, development of new activities, review, and explanation of the homework for the next class.	Increase of teacher and student attendance.	
Textbooks for *Se Liga/Acelera* students provided by IAS.	Student's learning levels are mapped at the beginning o the school year.	Monthly monitoring of proficiency.	Increase in the number of books read.	
Methodological proposal for the program's monitoring provided by IAS.	Once a week, mediators attend the classes of the students they are in charge of.	Monitoring of student's attendance and completion of the homework.	Increase in the % of homework completed.	
	Mediators provide feedback to teachers.	Mediators support teacher's pedagogic planning.		
	Monthly pedagogic meetings with teachers, mediators, and technical staff of the Secretary of Education.	Monthly pedagogical planning involving teachers, mediators, and technical staff of the Secretary of Education.		

Note: Authors' elaboration.

2.4 Data

In Brazil, the public education system is decentralized and the 26 states, the Federal district, and the 5,570 municipalities share the responsibilities for the provision of education. According to the 1988 Constitution, the municipal governments should give priority to early childhood and primary (grades 1 to 5) and lower secondary education (grades 6 to 9), and the state authorities to primary, lower, and upper secondary education (grades 10 to 12). Hence, the schools located in all the Brazilian municipalities can be managed either by the state government or can be under the management of the municipal government. The first group of schools is state-managed, and the second group of schools is locally-managed. Recife is one of the 184 municipalities of the state of Pernambuco. According to the 2020 Census of Education, the municipality of Recife has 205 primary schools managed by the municipal government, and two primary schools managed by the state government. The Municipal Department of Education is in charge of overseeing educational policies implemented in locally-managed schools, such as *Acelera*.

Since 1995, all private and public schools offering primary, lower, and upper secondary education participate in the annual Census of Education.[14] The Census is implemented by the National Institute of Educational Studies and Research (INEP), a research agency under the Brazilian Ministry of Education.[15] The Census collects information on (i) school facilities, such as libraries, sports courts, and science and computer labs; (ii) school infrastructure, such as filtered water, electricity, and internet access; (iii) social services, for example, school transportation and provision of meals; (iv) students, such as sex, the color of the skin, age, physical disabilities or mental illness, grade level, instruction time per day, class-size, subjects they are enrolled in, grade promotion, repetition and dropout rates; and (v) teachers, such as educational attainment, age, physical disabilities, subjects taught, and classes they are in charge of.

In 2000, the Department of Education of the state of Pernambuco designed an Education Assessment System (SAEPE) to monitor the performance and grade promotion of primary and secondary students.[16] Proficiency tests in Portuguese and math are annually applied to students of state and locally-managed schools across the 184 municipalities of the state. Up to

[14] Schools offering early childhood education also participate.

[15] INEP stands for *Instituto Nacional de Estudos e Pesquisas Educacionais.*

[16] SAEPE stands for *Sistema de Avaliação da Educação Básica de Pernambuco.*

2015, third and fifth graders were the primary education students included in the exam. Starting in 2016, second and fifth graders are the ones that answer the test. I normalize students' proficiency so that the treatment effect estimates can be measured in terms of standard deviations (SD). The use of the standardized measure prevents the analysis from being scale-sensitive and allows for comparability between grades and with other studies. Children also answer a socioeconomic questionnaire with information on their household infrastructure; parents' educational attainment; incentives from the family to continue studying; whether the teacher corrects their homework, if they already dropped out or repeated a grade; and if they did kindergarten. The SAEPE dataset is available at the student level.

The Information Technology company of the municipality of Recife (EMPREL) is in charge of collecting information on locally-managed schools.[17] EMPREL organizes annual data at the student level containing information on (i) the school and grade of enrollment; (ii) code of classroom, making it possible to identify student's peers; (iii) type of enrollment (*Acelera* or regular education) ; (iv) date of birth, from which one can calculate the difference between their age and the adequate age for the grade they are enrolled in; (v) sex; (vi) whether they attend school in the morning, afternoon or both; and (vii) their status at the end of the school year (promoted to the next grade, retained, or dropped out). Since each student has an enrollment code, it is possible to follow them over the years. The variable promoted to the next grade assumes values of 1 if the student is promoted to the next grade at the end of the school year, and 0 otherwise. The variable retained assumes values of 1 if the student is retained in the same grade at the end of the school year, and 0 otherwise. The variable dropped assumes values of 1 if the student dropout at the end of the school year, and 0 otherwise.

To assess the effects of *Acelera* on students' proficiency, I have to merge the EMPREL and SAEPE datasets. However, these datasets do not have the same students' identification numbers. Therefore, to conduct the merging, and then identify the proficiency of the students enrolled in locally-managed schools, I rely on students' names, dates of birth, grade, and school of enrollment. First, I perform the matching based on identifying the students with the exact same characters in both datasets. Second, to accommodate eventual typos in students' names not corrected in the data cleaning, then increasing the percentage of matched students,

[17] EMPREL stands for *Empresa Municipal de Informática.*

I incorporate a technique that performs the merging based on similar text patterns. I am able to find 70% of the students included in the proficiency test in the EMPREL dataset.[18] Not all the students enrolled in the locally-managed schools are included in SAEPE. There is the possibility that the school where they are enrolled is not included in the test (due to the small number of students enrolled in the grades assessed) or that the student does not go to school on the day of the assessment. Table B.2 shows the percentage of students that have proficiency scores. For example, on average 54.6% of third-graders have performance data available.[19]

By 2018, more than eight thousand students have participated in *Acelera*, and the number of schools offering the intervention significantly increases over time. In 2010, only 12 schools (5.7%) offer the program. In contrast, in 2018, the number jumps to 94, representing more than 40% of schools in the municipality (Table 2.2). I observe that, on average, *Acelera* schools are bigger, in terms of the number of students, have more classrooms available to create additional classes to offer the intervention, and, as expected, have a higher percentage of students with age-grade distortion and lower proficiency scores (Table B.3).

At the beginning of the school year, once the Department of Education selects the schools to implement *Acelera*, I observe that school boards mainly select fourth and fifth graders, as 85.4% of the treated students are enrolled in these grades. Among the participating students, 86.3% are at least one year older than the adequate age for their grade. Since I do not have access to the results of the proficiency assessment applied at the beginning of the school year, which is also used to identify potential program participants, I define the eligible group as the students that have at least one year of age-grade distortion (as 86.3% of program participants attend this criterion).[20]

Among the schools that offer *Acelera*, 24.6% of the students students are included in the intervention (Table 2.2), creating a significant overlap of program participants and eligible students enrolled in regular education (Figure B.1). I observe that the way the program is offered creates two groups to which *Acelera* students can be compared. The first comparison

[18] 30% of the students included in SAEPE are not found in the EMPREL dataset because they do not follow the criteria established for the merging.

[19] The remaining 45.5% either did not go to school when the test was applied or these students were not found in the SAEPE dataset because they did not attend the criteria for the merging.

[20] The municipal government of Recife only shares the EMPREL and SAEPE datasets, which do not contain data on the tests applied at the beginning of the school year. Each school applies a distinct test and they do not set up an organized dataset that would allow me to check students' performance.

group is formed by eligible students enrolled in schools offering *Acelera* but that are not included in the intervention. The second comparison group is formed by eligible students in the remaining locally-managed schools in Recife, the ones that do not offer *Acelera*.

In Recife, between 2009, one year before *Acelera*, and 2018, the municipality experienced a decrease of 70% in dropout rates and an increase in the proficiency of fifth graders (Figure B.4). To assess whether *Acelera* is associated with the improvement of these educational indicators, I compare the program participants, the *treatment group*, with students that, although eligible to participate, are not included in the intervention, the *comparison group*.

Table 2.2: Sample of the study, first to fifth grade (2010-2018)

	Schools offering only Regular Education *Students with at least one year of age-grade distortion*		**Schools offering Acelera**				**Acelera Enrollment by grade**				
	N. schools	Enrollment	N. schools	Enrollment by type of education: Regular classes	Enrollment by type of education: Acelera classes	Enrollments in Acelera with at least one year of age-grade distortion	First grade	Second grade	Third grade	Fourth grade	Fifth grade
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
2010	200	8,954	12	494	251	222	2	3	15	63	168
as %			5.7	69.0	35.1	88.4	0.8	1.2	6.0	25.1	66.9
2011	166	6,300	45	1,496	667	535	6	7	59	267	328
as %			21.3	73.7	32.8	80.2	0.9	1.0	8.8	40.0	49.2
2012	149	5,621	62	2,520	947	758	6	18	85	355	483
as %			29.4	76.9	28.9	80.0	0.6	1.9	9.0	37.5	51.0
2013	171	6,431	52	2,483	956	783	27	10	93	436	390
as %			23.3	76.0	29.3	81.9	2.8	1.0	9.7	45.6	40.8
2014	193	7,860	31	1,673	552	460	1	10	54	274	213
as %			13.8	78.4	25.9	83.3	0.2	1.8	9.8	49.6	38.6
2015	168	7,330	51	3,001	850	729	1	24	124	346	355
as %			23.3	80.5	22.8	85.8	0.1	2.8	14.6	40.7	41.8
2016	146	5,955	73	4,755	1,117	989	0	33	101	443	540
as %			33.3	82.8	19.4	88.5	0.0	3.3	10.2	44.8	54.6
2017	132	5,086	86	5,218	1,455	1,310	1	52	206	586	610
as %			39.4	79.9	22.3	90.0	0.1	3.6	14.2	40.3	41.9
2018	124	4,424	94	5,358	1,628	1,483	5	42	243	659	679
as %			43.1	78.3	23.8	91.1	0.3	2.6	14.9	40.5	41.7
Total		57,961		26,998	8,423	7,269	49	199	980	3,429	3,766
as %					24.6	86.3	0.6	2.4	11.6	40.7	44.7

Notes: Column I: number of schools that only offer Regular Education. Columns II and IV: enrollment of primary students with at least one year of age-grade distortion in regular education. Column III: number of schools that offer regular education and *Acelera*. Column V: enrollment of primary students in *Acelera* . Column VI: enrollment of primary students with at least one year of age-grade distortion in *Acelera*. Columns VII to X: enrollment of *Acelera* students by grade. Source: EMPREL.

To assess the impact of *Acelera* on grade promotion, age distortion, and dropout, I work with a pooled sample of first to fifth graders from 2008 to 2014.[21] The data allow that since these educational indicators are available for the panel of students, regardless of their grade of enrollment.

However, to investigate the impacts of the intervention on students' proficiency in reading and math, I need to focus on fourth graders that are either enrolled in *Acelera* or are eligible to. This happens because the proficiency score in the standardized exam (SAEPE) is not available for all primary education grades. Until 2015, only third and fifth graders were subject to the exam.[22] Therefore, by working on a sample of fourth graders, I can compare their proficiency before the intervention (when they are in third grade) and after that (when they are in fifth grade). In this sense, I restrict the sample to fourth-graders whose proficiency data is available for both third and fifth grades (Table 2.3).

In the comparison group, approximately 34.7% follow these criteria and among *Acelera* participants the percentage is 9.6%. The low percentages are due to a combination of factors. First, 30% of the students included in SAEPE are not found in the EMPREL dataset, mostly because of typos in their names, not possible to be corrected with the data cleaning. Second, not all the students participate in the proficiency assessment applied by the state government. Third, for 27% of fourth-graders: i) the fourth grade is their first grade of enrollment in a locally-managed school (they were probably previously enrolled in the private or state networks); and ii) the fourth grade is their last grade of enrollment in a locally-managed school (they either dropped out or migrate to the private or state networks). Since I only have the enrollment data for locally-managed schools, I cannot find data on these students for their third and fifth-grade enrollment. Fourth, as the ultimate goal of *Acelera* is to make students jump up to two grades, fourth graders can jump to the sixth grade, as shown in Figure B.3. In fact, this is the case for 42% of them. Therefore, students in this situation do not participate in the standardized proficiency assessment applied to fifth graders.

One may wonder whether among *Acelera* participants there is a selection of the best performers to do the standardized proficiency test. I find evidence that this is not the case. When I compare

[21] In 2015 and 2016, 12.3% and 17.6% of students do not have data for grade promotion, repetition, or dropout rates, mostly students enrolled in *Acelera*. Therefore, I focus the analysis on the years that we have reliable information for the status of the students at the end of the school year.

[22] Starting in 2016, second and fifth graders are the students included in the assessment.

the students that do the test and the ones that do not, there are no significant differences between grade promotion and repetition rates. Besides that, the past proficiency of absent students is significantly higher, suggesting that the students that do the test are actually the low performers (Table B.4). This might be because the best performers are the ones that jump to sixth grade after participating in *Acelera* in the fourth grade. Therefore, since the sample of treated students is mostly fourth graders with lower past performances, my estimates potentially underestimate the true impact of the intervention.

Table 2.3: Sample of fourth graders, (2011-2017)

	Regular schools				**Schools offering Acelera**								
	Students with at least one year of age-grade distortion				Students enrolled in Acelera					*Students enrolled in Regular Education with at least one year of age-grade distortion*			
	Enrollments	% of students with proficiency in:			Enrollments	Jumped	% of students with proficiency in:			Enrollments	% of students with proficiency in:		
	4th grade	3rd grade	5th grade	3rd and 5th grade	4th grade	to 6th grade	3rd grade	5th grade	3rd and 5th grade	4th grade	3rd grade	5th grade	3rd and 5th grade
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XII
2011	1,187	46.7	55.9	31.8	267	30.3	35.2	12.7	5.2	231	38.1	54.1	24.2
2012	1,227	61.3	49.6	34.7	355	24.5	51.0	16.3	9.6	361	54.3	47.6	29.4
2013	1,269	63.0	62.0	41.4	434	22.1	64.5	15.4	9.2	384	65.1	58.1	42.7
2014	1,238	44.6	54.7	27.6	274	14.2	56.9	23.0	14.6	216	46.8	65.3	35.2
2015	1,325	55.1	60.4	35.1	344	14.2	63.4	20.9	12.2	451	50.8	58.1	32.2
2016	1,164	65.3	75.9	53.3	435	16.6	65.5	22.3	15.4	657	59.5	73.1	46.6
2017	1,007	36.2	72.1	27.8	584	18.2	44.9	15.9	7.7	774	31.4	62.0	21.6
Total	8,417	55.1	60.4	34.7	2,693	18.2	56.9	16.3	9.6	3,074	50.8	58.1	32.2

Source: EMPREL and SAEPE.

2.5 Empirical Strategy

To assess the impact of *Acelera* on students' proficiency in Portuguese and math, grade promotion, age-grade distortion, and dropout, I explore the main criteria to be included in the intervention. I define as *eligible group* the students that are at least one year older than the adequate age for their grade of enrollment. I then compare the program participants, the *treatment group*, with eligible students not included in the intervention, *comparison group* (Table 2.2).

2.5.1 Standard difference-in-differences

I first employ a standard DiD specification on a sample of eligible students to estimate the policy impacts:

$$y_{igst} = \alpha_0 + \gamma T_{igst} + \alpha_1 X'_{igst} + \alpha_2 Z'_{gst} + \alpha_3 W'_{st} + \rho_i + \phi_g + \mu_s + \theta_t + \upsilon_{igst} \tag{2.1}$$

In which y_{igst} is one of the dependent variables of the study of the student i, in grade g, in school s, in year t. T_{igst} is a dummy equal to 1 for the year the student i is included in *Acelera* and in the years after that, and 0 otherwise. Aiming to increase precision and account for potential time-variant confounders, I include a set of controls at the student, grade, and school levels. X'_{igst} are the controls at the student level (sex and difference between students' age and the adequate age for their grade). Z'_{gst} are the controls at the grade level (students per class and difference in years between the youngest and the oldest student of the class). W'_{st} are the controls at the school level (library, computer, science lab, internet access, sports court, number of employees, and access to energy and water supply). ρ_i, ϕ_g, μ_s, θ_t are fixed effects for students, grade, school and year, respectively. The parameter of interest, γ, is the average treatment effect on the treated (ATT), that is, the average effect of *Acelera* on its participants. The standard error is clustered at the student level.

On the one hand, since not all the schools of Recife offer *Acelera*, my approach allows me to compare similar students with the difference that one is enrolled in a school that offers the intervention and the other student is not. Although *Acelera* guidelines specify that an eligible student can change schools to participate in the program, the data indicate that this is only the case for 20% of them. Therefore, there are several students eligible for the program but that are not enrolled in a *Acelera* school. On the other hand, since I observe significant differences between regular and *Acelera* schools, one may wonder whether treatment and comparison groups are affected differently (Table B.3). To account for this, I also run the equation 2.1 restricting the analysis to treated and comparison students of the schools that offer *Acelera* in at least one year between 2010 and 2018.

Another identification threat are the non-observable factors that lead to the selection of *Acelera* participants. Besides the age-grade distortion, the school board selects participant students based on their scores on proficiency tests applied at the beginning of the school year, which I do not have access to.[23] Also, not all selected students end up participating if their parents do not authorize it. To account for these factors, besides controlling for students' fixed effects in equation 2.1, I employ a propensity score matching using the sample of eligible students. For each year t between 2010 and 2014 and grade g, I run the following specification to match *Acelera* participants to comparison students:

$$\Pi[Acelera_{igst} = 1/X] = \lambda X'_{igst} + \mu_s + \theta_t + \epsilon_{igst} \tag{2.2}$$

in which $Acelera_{igst}$ is equal to 1 if the student i is included in *Acelera* in t, and 0 otherwise. X'_{igst} are the controls at the student level, such as the level of age-grade distortion (the difference in years between students' age and the adequate age of the grade they are enrolled in), whether they have already participated of the *Se Liga* intervention and their status in $t-1$ (jumped to the next grade or retained). α_s and θ_t are school and year fixed effects, respectively.[24] I then append the sample of matched students, only the ones in the common support, and then run equation 2.1. Figure B.5 shows that the probability of being treated is very similar for treatment and comparison students.

[23] The municipal government of Recife only shares the Emprel and SAEPE datasets, which do not contain data on the tests applied at the beginning of the school year. Each school uses a distinct test and they do not set up an organized dataset that would allow me to check students' performance.

[24] *Se Liga* also targets students with at least one year of age-grade distortion but focuses on the ones that do not know how to read and write.

2.5.2 Leads and lags

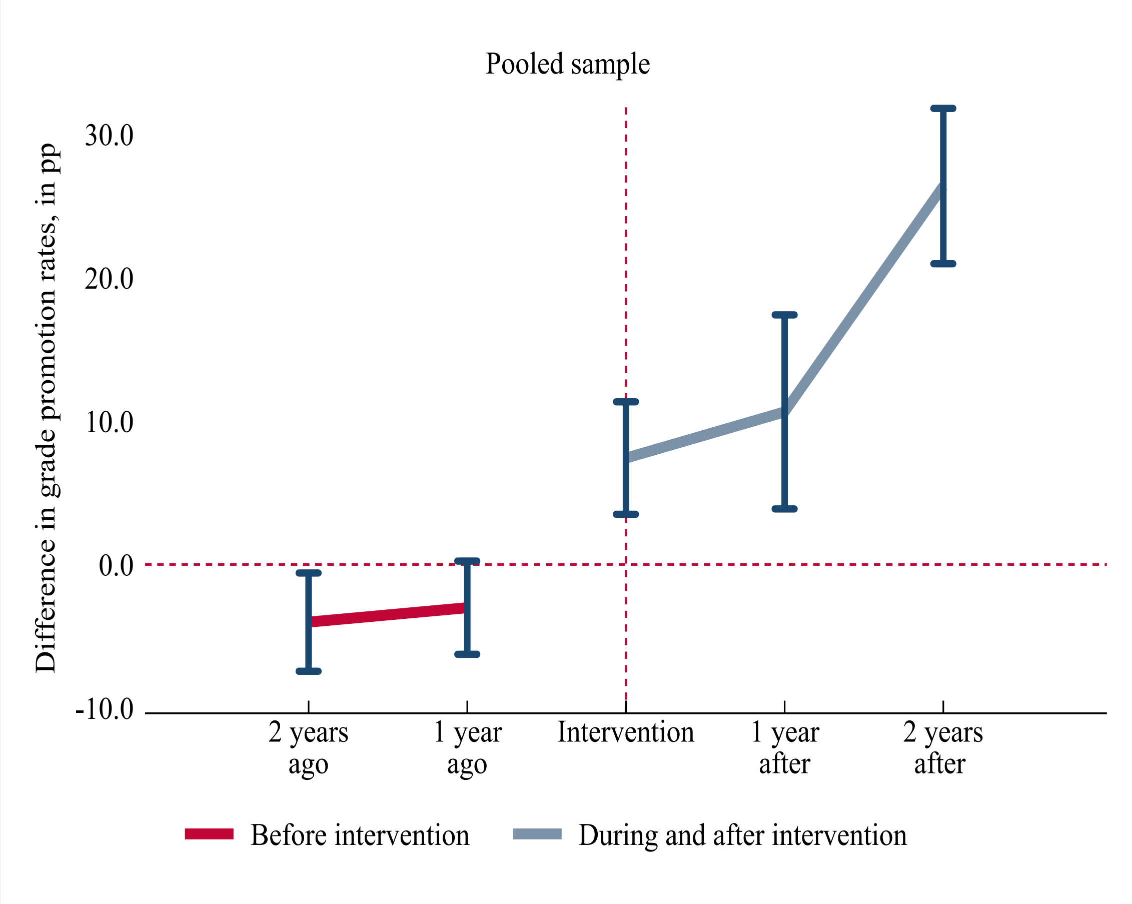
In the second approach to investigate the causal impact of *Acelera*, I explore the fact that the treatment status changes at different times. In the difference in differences methodology, the students are considered treated in the year they receive the intervention and after that, (when they return to regular education). For the dependent variables grade promotion, age distortion, and dropout, the structure of the data allows me to enrich the analysis to check whether the treatment status predicts the educational outcomes and not the educational outcomes predict treatment status. Consider D_{it} a dummy equal to 1 if the student i is included in *Acelera* in year t, and 0 otherwise. Assume I observe the students for m years after the treatment and q years before it. Angrist and Pischke (2008) suggest the following specification:

$$y_{igst} = \alpha_0 + \sum_{\tau=1}^{m} \beta_{-\tau} D_{i,t-\tau} + \sum_{\tau=1}^{q} \beta_{+\tau} D_{i,t+\tau} + \alpha_1 X'_{igst} + \alpha_2 Z'_{gst} + \alpha_3 W'_{st} + \rho_i + \phi_g + \mu_s + \theta_t + \upsilon_{igst} \quad (2.3)$$

In which $D_{i,t-\tau}$ and $D_{i,t+\tau}$ assume value 1 in τ periods after/before the treatment, respectively, and 0 otherwise. Therefore, the specification allows for m lags (β_{-1},β_{-2},....,β_{-m}) or post-treatment effects and q leads (β_{+1},β_{+2},....,β_{+q}) or anticipatory effects. The lags make it possible to check whether the effects of *Acelera* grow or fade as time passes; and if one believes that $D_{i,t}$ causes Y_{igst} and not vice versa, the leads should not matter in equation 2.3. As with the standard difference-in-differences approach, I run the leads and lags framework on the overall sample and on the matched student sample.

I observe that there are no significant differences in grade promotion rates of participants and non-participants immediately prior to the intervention, suggesting that, as expected, it is not this outcome that predicts treatment (Figure 2.1).

Figure 2.1: Leads and lags estimates for grade promotion, pooled sample (2008-2014)



Note: Author's estimate of the equation 2.3. A pooled sample of first to fifth graders. The estimates consider the sample of schools that offer *Acelera* in at least one year between 2010 and 2014. Source: EMPREL.

2.6 Results

Table 2.4 presents the results of the difference-in-differences and leads and lags framework for grade promotion, dropout, and age-grade distortion considering the pooled sample of first to fifth-graders. The DiD approach provides the average treatment effect of *Acelera* for the year of the treatment and after that (when they return to regular education).[25] The leads and lags approach makes it possible to disentangle the average impact of the intervention for the year the student is treated and for each one of the years after that, allowing us to check if the impacts grow or fade within time (Figure 2.1).

The baseline DiD specification suggests that *Acelera* led to an 16 percentage points (pp) increase in grade promotion of primary education students (column I of Table 2.4). Since the

[25]The treatment status under the DiD approach is equal to zero before the student joins the intervention, and equal to one in the year the student is allocated into an *Acelera* classroom and also in the years after that.

average grade promotion of participants before the intervention is 70.7%, the 16 pp estimate is equivalent to a jump of 22.6%. The baseline leads and lags estimate indicates a 4.5 pp increase in grade promotion rates in the year of participation, equivalent to a 6.3% increase (column II of Table 2.4). The fact that DiD estimates are bigger than the leads and lags suggests that the program increases grade promotion in the year students are treated and continue impacting them after the return to regular education.

One of the concerns of the baseline specification is that, although I compare only eligible to *Acelera* students, part of the comparison group is from schools that do not offer the intervention. Table B.3 shows that these schools are significantly different from the ones that offer *Acelera*. In this sense, the estimates might be biased if non-observable time-varying factors of participant schools are correlated with both the allocation of a set of their students into *Acelera* classes and the dependent variables of the study. To test if this is the case, I run the analysis on a sample that includes in the comparison group only eligible students from schools that offer *Acelera* (Columns III and IV of Table 2.4). I observe that the estimates are similar to the ones presented in columns I and II.[26]

Another identification threat is that I do not observe all the factors that determine the selection of a student into a *Acelera* class. I then take advantage of the fact that schools do not include all the students with age-grade distortion in the intervention, creating an overlap in the distribution of the years of age-grade distortion between participants and eligible non-participants (Figure B.1). I explore this by matching students based on their characteristics to compare the most similar groups, with the difference that some are selected to participate and others not (columns V to VIII of Table 2.4). The estimates are similar, although a bit bigger than the non-matched sample. The differences might suggest that the analysis with the non-matched sample underestimates the program's impact since the comparison group includes students that are more different than the participants and probably have better educational indicators that lead them not to be included in *Acelera*.

Students have three possible outcomes at the end of the school year: jump to the next grade, be retained, or drop out. I find evidence of a significant increase in grade promotion, which means a decrease in retention and possibly a decrease in dropout. The results shown in Table 2.4

[26] Although the point estimate of leads and lags shown in column II is smaller than the one in column IV, it is possible to see that estimate (4.5 pp) is included in confidence interval for the column IV estimates.

show that the intervention does not decrease dropout rates. Taking into account that *Acelera* participants are the most vulnerable students and, therefore, would be more likely to drop out, this result can be seen as a positive outcome. It is the lower bound estimate and indicates that participant students are as likely as non-participants to drop out. This might be associated with the dynamics of *Acelera* classes as students are constantly motivated and also to the school curriculum, which is more adequate to their learning level and makes it easier to keep pace with the classes.

In addition to that, as a consequence of the increase in grade promotion, I observe a decrease in age-grade distortion. The DiD baseline estimate suggests a decrease in the percentage of students with at least one year of age-grade distortion of 10 pp, equivalent to a decrease of 17%.

Table 2.4: Impact of *Acelera* on grade promotion, dropout, and age-grade distortion, first-fifth grade (2008-2014)

	No matching				Matching			
	All schools		Only Acelera schools		All schools		Only Acelera schools	
	DiD	Leads, lags	DiD	Leads, lags	DiD	Leads, lags	DiD	Leads, lags
	I	II	III	IV	I	II	III	IV
	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95
Grade-promotion								
ATT	15.99***	4.45***	15.01***	7.75***	18.75***	7.28***	15.21***	7.55***
	(1.31)	(1.69)	(1.45)	(1.97)	(1.44)	(1.91)	(1.67)	(2.23)
	[13.42,18.56]	[1.14,7.77]	[12.16,17.86]	[3.90,11.61]	[15.93,21.57]	[3.54,11.03]	[11.93,18.48]	[3.17,11.93]
Obs	77628	77628	26486	26486	37157	37157	21789	21789
R2	0.31	0.32	0.3	0.31	0.29	0.29	0.29	0.3
Number of schools	224	224	216	216	220	220	212	212
Treatment Group								
Num. students	1657	1657	1651	1651	1628	1628	1623	1623
Mean outcome	70.73	70.73	70.7	70.7	70.83	70.83	70.82	70.82
SD	45.5	45.5	45.52	45.52	45.46	45.46	45.46	45.46
ATT in sd	0.35	0.1	0.33	0.17	0.41	0.16	0.33	0.17
Comparison Group								
Num. students	17118	17118	4429	4429	6482	6482	3110	3110
Mean outcome	74.86	74.86	72.99	72.99	74.96	74.96	73.97	73.97
Dropout								
ATT	0.63	-0.75	0.98**	-0.57	0.54	-1.08*	0.61	-1.33
	(0.44)	(0.58)	(0.49)	(0.65)	(0.49)	(0.65)	(0.64)	(0.84)
	[-0.24,1.49]	[-1.90,0.39]	[0.01,1.94]	[-1.84,0.71]	[-0.41,1.50]	[-2.37,0.20]	[-0.65,1.87]	[-2.98,0.32]
Obs	77628	77628	26486	26486	37157	37157	21789	21789
R2	0.02	0.02	0.04	0.04	0.04	0.04	0.05	0.05
Treatment Group								
Mean outcome	1	1	1	1	0.78	0.78	0.78	0.78
SD	9.95	9.95	9.97	9.97	8.81	8.81	8.82	8.82
ATT in sd	0.06	-0.08	0.1	-0.06	0.06	-0.12	0.07	-0.15
Comparison Group								
Mean outcome	1.33	1.33	1.17	1.17	1.15	1.15	1.42	1.42
Age-grade distortion								
ATT	-10.32***	-6.66***	-13.00***	-6.45***	-8.92***	-1.15	-10.74***	-1.45
	(0.92)	(1.38)	(1.07)	(1.58)	(1.03)	(1.50)	(1.18)	(1.71)
	[-12.12,-8.53]	[-9.37,-3.95]	[-15.11,-10.90]	[-9.54,-3.36]	[-10.94,-6.89]	[-4.09,1.78]	[-13.05,-8.42]	[-4.79,1.90]
Obs	78190	78190	26727	26727	37399	37399	21958	21958
R2	0.69	0.69	0.69	0.69	0.65	0.65	0.66	0.66
Treatment Group								
Mean outcome	60.16	60.16	60.14	60.14	60.07	60.07	60.03	60.03
SD	48.96	48.96	48.97	48.97	48.98	48.98	48.99	48.99
ATT in sd	-0.21	-0.14	-0.27	-0.13	-0.18	-0.02	-0.22	-0.03
Comparison Group								
Mean outcome	67.57	67.57	64.6	64.6	70.48	70.48	68.92	68.92

Notes: ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Equations 2.1 and 2.3. Standard errors in parenthesis and clustered at the student level. Square brackets show a 95% confidence interval. The columns "no matching" include the sample of all eligible students. The columns "matching" include only matched students from treatment and comparison groups. The columns "All schools" include all locally-managed schools in Recife. The columns "only *Acelera* schools" include only students that have ever been eligible for treatment in a school that offer *Acelera*. ATT regressions coefficients are converted to percentages (multiplied by 100).

To investigate the underlying channels, I propose an additional specification in which I interact the treatment with some of the students' characteristics, such as sex, the difference between their age and the adequate age of the grade they are enrolled in, and the age difference between the youngest and the oldest student of their classroom (a measure of classroom heterogeneity) (Table 2.5).

First, I observe that the program does not seem to affect boys differently than girls.[27] Second, since *Acelera* classes gather students with distinct levels of age-grade distortion, one may wonder whether the students with higher levels of delay could face more challenges during the project implementation. The higher levels of distortion, besides indicating that these students are among the low performers, might also be linked to a lack of motivation and self-esteem problems, as these students have watched their peers change grades, while they are retained. Indeed, the results indicate that the higher the age distortion, the lower the jump in grade promotion rates due to *Acelera*. Third, as *Acelera* classes are set up with students from different grades, which could raise questions on whether the higher difference between the youngest and the oldest of the class would impact the program's implementation. I do not find evidence that this is the case.

[27]The variable *sex* is equal to 1 for boys, and 0 for girls.

Table 2.5: Heterogeneity of the impact of *Acelera* on grade promotion, dropout, and age-grade distortion, first-fifth grade (2008-2014)

	No matching						Matching					
	All schools			Only Acelera schools			All schools			Only Acelera schools		
	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD
	I	II	III	IV	V	VI	I	II	III	IV	V	VI
	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95	b/se/ci95
Grade-promotion												
ATT	14.12***	5.38*	68.56***	13.53***	11.12***	66.08***	17.22***	13.28***	67.41***	13.90***	13.23***	62.94***
	(1.59)	(2.77)	(2.67)	(1.70)	(3.19)	(2.79)	(1.68)	(3.10)	(2.69)	(1.89)	(3.53)	(3.04)
ATT versus sex	4.75*			3.69			3.85*			3.26		
	(2.43)			(2.28)			(2.25)			(2.24)		
ATT versus age dif class		3.98***			1.42			1.95*			0.7	
		(1.03)			(1.10)			(1.07)			(1.18)	
ATT versus distortion			-26.20***			-24.97***			-24.54***			-23.82***
			(1.28)			(1.29)			(1.28)			(1.34)
Obs	77628	77628	77628	26486	26486	26486	37157	37157	37157	21789	21789	21789
R2	0.31	0.31	0.32	0.3	0.3	0.32	0.29	0.29	0.31	0.29	0.29	0.32
Dropout												
ATT	0.86	-1.79*	-4.69***	1.20**	-1.53	-5.40***	0.56	-2.06*	-5.48***	0.59	-1.71	-6.12***
	(0.55)	(1.02)	(1.13)	(0.59)	(1.13)	(1.22)	(0.58)	(1.11)	(1.17)	(0.72)	(1.34)	(1.34)
ATT versus sex	-0.59			-0.55			-0.03			0.04		
	(0.83)			(0.82)			(0.80)			(0.79)		
ATT versus age dif class		0.91**			0.91**			0.93**			0.82*	
		(0.39)			(0.43)			(0.42)			(0.47)	
ATT versus distortion			2.65***			3.12***			3.04***			3.36***
			(0.65)			(0.67)			(0.66)			(0.68)
Obs	77628	77628	77628	26486	26486	26486	37157	37157	37157	21789	21789	21789
R2	0.02	0.02	0.03	0.04	0.04	0.05	0.04	0.04	0.05	0.05	0.05	0.06

Notes: ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Equations 2.1 and 2.3. Standard errors in parenthesis and clustered at the student level. The columns "no matching" include the sample of all eligible students. The columns "matching" include only matched students from treatment and comparison groups. The columns "All schools" include all locally-managed schools in Recife. The columns "only *Acelera* schools" include only students that have ever been eligible for treatment in a school that offer *Acelera*. ATT regressions coefficients are converted to percentages (multiplied by 100).

Table 2.6 presents the results of the difference-in-differences approach for students' proficiency in Portuguese and math. I do not find evidence that *Acelera* leads to an increase in learning levels. However, this result might be interpreted with caution. As described in subsection 2.4, I estimate the impact of the intervention on *Acelera* participants that join the intervention in the fourth grade. Since the ultimate goal of *Acelera* is to make students jump up to two grades, boosting the decrease in age-grade distortion, fourth graders with proficiency scores in the fifth grade are the ones that are not promoted to the six-grade and probability might be the low-performers of the class. Table B.4 shows that 50% of fourth graders that do not have proficiency scores in the fifth-grade jump to sixth grade after participating in *Acelera*. Also, their pre-treatment proficiency is higher than the one observed by fourth-graders that have post-treatment proficiency scores. Therefore, I compare eligible non-participant students with treated students that are among the low performers in the class. I then end up underestimating the true impact of the intervention. Also, this result indicates that participant students promoted to the next grade are not learning less than the comparison group. If this was the case, it would be an indication that teachers promote participant students to the next grade even though they keep learning less than the comparison group.

Table 2.6: Impact of *Acelera* on students' performance in reading and math, fourth grade (2008-2018)

	Portuguese				Math			
	No matching		Matching		No matching		Matching	
	All schools	Acelera schools	All schools	Acelera schools	All schools	Acelera schools	All schools	Acelera schools
	I	II	III	IV	I	II	III	IV
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
ATT	-0.01	0	0.03	0.03	-0.02*	-0.01	0.03	0.03
se	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
95% CI	[-0.04,0.01]	[-0.04,0.03]	[-0.01,0.08]	[-0.02,0.08]	[-0.05,0.00]	[-0.04,0.02]	[-0.01,0.07]	[-0.02,0.08]
Obs	8620	2618	808	632	7488	2416	734	632
R2	0.19	0.26	0.34	0.35	0.11	0.2	0.26	0.35
Treatment Group	254	254	133	131	233	233	120	131
Acelera students	449.4	449.4	451.71	453.56	456.26	456.26	452.65	453.56
Mean outcome, in %	89.55	89.55	87.37	88.06	86.8	86.8	79.01	88.06
Comparison Group								
Students in regular education	4056	1055	271	185	3511	975	247	185
Mean outcome, in %	466.54	468.65	471.15	466.32	464.82	466.15	464.11	466.32
Num. schools	212	169	107	92	212	163	100	92

Notes: ***, **, and * indicate significance at the 1, 5, and 10% critical level. Equation 2.1. Equations 2.1 and 2.3. Standard errors in parenthesis and clustered at the student level. Square brackets show a 95% confidence interval. The columns "no matching" include the sample of all eligible students. The columns "matching" include only matched students from treatment and comparison groups. The columns "All schools" include all locally-managed schools in Recife. The columns "Acelera schools" include only students from schools that offered *Acelera* in at least one year between 2010 and 2018. The dependent variables (students' proficiency) are normalized.

2.7 Discussion and Conclusion

In the last decades, Brazil has made huge progress in universalizing access to primary education, such that, the net enrollment rate of children from 6 to 14 years old reaches 99.3%, more than the average observed for developing countries. Nonetheless, more than one out of five children have not yet finished lower secondary education by age 16, the age that they are already supposed to be in high school.[28] Also, almost 40% of fifth-graders do not have an adequate level of proficiency in Portuguese, and the percentage is close to 50% in math.[29] Those lagging behind will struggle in later grades and might never catch up, leading to grade retention, dropout and, as a consequence, an increase in age-grade distortion. In this context, interventions aimed to adequate the school content to students' level of education, *Teaching at the Right Level*, can be an important tool.

Since 2010, the Department of Education of Recife, in the state of Pernambuco, has been implementing *Acelera*, a program that groups primary education students lagging behind their peers and who are at least one year older than the adequate age for their grade. The students are then allocated to a new class for the whole school year. The intervention aims to increase learning levels, grade promotion, and decrease dropout and age-grade distortion.

In this chapter, I employ a difference-in-differences analysis on a rich dataset at the student level to assess the impact of the program from 2010 to 2018. To estimate the impact of *Acelera* on grade promotion, dropout, and age-grade distortion, the data available allows me to follow a panel of first to fifth-graders. I find evidence that the program not only increases grade promotion in the year of participation but also in the years after that when the participant students return to regular education. The results show an increase in grade promotion rates of 16 percentage points, a jump of 22.6% when compared to baseline levels. Also, my estimates suggest a decrease of 17% in the percentage of students with age-grade distortion. The heterogeneity analysis indicates that the higher the levels of age-grade distortion, the less the students benefit from the intervention. The sooner the eligible students are included in the intervention, the better.

[28] See the Educational Report of *Todos pela Educação* in this link: https://todospelaeducacao.org.br/wordpress/wp-content/uploads/2020/10/Anuario-Brasileiro-Educacao-Basica-2020-web-outubro.pdf

[29] See the Educational Report of *Todos pela Educação* in this link: https://todospelaeducacao.org.br/wordpress/wp-content/uploads/2021/09/relatorio-de-aprendizagem.pdf

Nonetheless, one may wonder whether the students are promoted to the next grade because they actually achieve an adequate level of proficiency or as a simple consequence of the program participation. I then investigate the effects of *Acelera* on students' proficiency in Portuguese and math. I do not find evidence that *Acelera* increases students' learning levels. However, this result might be interpreted with caution as it is a lower-bound estimate of the true impact of the intervention. This estimate suggests that, at least, participant students promoted to the next grade are not learning less than the comparison group.

Overall, the interventions aimed at adequate school content with children's level of education seem an important tool to increase grade promotion and decrease age distortion.

3 Estimating the efficiency of public primary education expenditures in Brazil

3.1 Introduction

Education has a crucial role in economic activity across the world. The more a country's workforce cognitive skills, the higher its economic growth and per capita income. Family background, individual abilities, and formal schooling determine those cognitive skills. The latter is the one the policymakers emphasize as it is the most directly affected by public policies (Hanushek (2006)). In this context, conventional wisdom states that public schools require additional resources to increase students' performance. However, the available evidence points out that overall schools fail to use their resources efficiently, suggesting that there is room to increase students' learning levels without additional expenditures. To estimate how efficiently the Brazilian municipalities use their resources, I employ a Data Envelopment Analysis to investigate the relationship between per-pupil expenditure and students' proficiency in reading and math, as well as grade promotion.

Data Envelopment Analysis is a non-parametric methodology to assess the efficiency of decision-making units (DMUs), such as municipalities, states, or countries. DEA consists in solving a linear programming model that compares the inputs being used, such as expenditure per student, students per class, pupil-teacher ratio, and the outputs achieved, such as students' proficiency, grade promotion, and high-school completion. The model can easily accommodate multiple outputs and inputs and it is not necessary to specify any functional form (Sutherland et al. (2007)). The model establishes a production frontier with DMUs that achieve the maximum output considering the inputs being used. The DMUs on the frontier are then used as a benchmark by the ones that do not reach the frontier.

Given the technology available, the output-oriented efficiency analysis measures the difference between what the DMU achieves and what would be possible to achieve if it operates on the production frontier, that is, the shortfall of outputs for a given level of inputs. The input-oriented efficiency analysis measures the difference between the inputs being used and the minimum amount that would be enough to reach the same output level if the DMU operates on the frontier, that is, the excess of input consumption (Mattos and Terra (2015a)).

In this Chapter, the DMUs are the municipal governments of the 5,570 Brazilian municipalities. The output under evaluation is the last data available on Brazil's most important primary education indicator, the 2019 Education Development Index (IDEB). The index combines students' proficiency in reading and math and grade promotion. Since the IDEB for primary education is a function of the investment made by the municipal governments in children from their first to fifth grade, my model has five inputs of per-pupil expenditure, one for each grade of primary education.

The group-frontier results indicate that Brazilian municipalities efficiently use between 72% to 83% of their educational resources. This suggests that if the local authorities could reach the production frontier, there would be a fiscal space of at least 86 billion BRL. An amount that could be allocated to interventions to increase students' performance in a post-pandemic context where they are so much needed. To have an idea of what 86 billion BRL represents in Brazil, it is more than twice the annual 2022 *Bolsa Família* budget, the most important conditional cash transfer in the country reaching more than 10 million families.[1] The results also suggest that municipalities that have between 5 and 50 thousand inhabitants are the ones that mostly make the meta-frontier.

According to the Education at a Glance (2021), the expenditure per primary education student in Brazil is approximately two and a half times lower than the OECD average. In that sense, it is important to point out that municipalities, even once they reach the efficiency frontier, still might require additional resources to improve their educational outcomes. It is possible that their current level of expenditure is not the optimal one required for students to achieve an adequate level of proficiency. In this case, governments need to increase per-pupil expenditure and expand their operation scale along the efficient frontier. The DEA allows me to identify examples of municipalities in this situation. Nova Canaã in Bahia, Parnaguá in Piauí, and Cruz do Espírito Santo in Paraíba are all on the frontier. The municipalities have an IDEB of 3.6, almost 40% lower than the target of 5.7 established by the Ministry of Education. However, compared to the other municipal governments in the analysis, these three municipalities achieve the maximum output level given their expenditure per student, which is nearly 35% lower than

[1] In 2022, the *Bolsa Família* was replaced by Auxílio Brasil.

the national average. They are examples of DMUs that achieve the frontier but need additional resources to improve their outcomes.

This chapter has several contributions to the literature. To my knowledge, this is the first Data Envelopment Analysis exploring the efficiency of education expenditures using the last data available on the most important educational indicator in Brazil (2019 IDEB). The analysis provides an estimate of the fiscal space that would be achieved if municipalities reach the production frontier. For each municipality under evaluation, the model provides a list of benchmark municipalities, that is, DMUs that are similar in terms of inputs being employed but that are able to reach the efficiency frontier. This would allow local authorities to share experiences and best practices. Increasing efficiency would provide resources that could be employed in education interventions in a country that was severely hit by the Covid-19 outbreak. In Brazil, most schools were closed for at least one year and students faced a significant decrease in performance. In addition, the method allows the ranking of the municipalities according to their efficiency. This measure could be used to reward local authorities that achieve a certain level of efficiency. Also, the combination of the following approaches consists of innovations in the efficiency literature. My estimates allow for distinct production function technologies, an important assumption considering a country as diverse as Brazil. Since family background plays an important role in students' performance, I include their mothers' education as a non-discretionary input. Figure C.1 shows that by not including mothers' education in the analysis, one underestimates the efficiency of the DMUs under evaluation.[2] I run super-efficiency models to detect outliers in order that I do not underestimate the efficiency of educational expenditures across the country. Finally, I also run bootstrap models to calculate confidence intervals for the efficiency estimates.

Apart from this introduction, this chapter is organized as follows: Section 3.2 presents the related literature. Section 3.3 introduces the data available to perform the analysis. Section 3.4 discuss the empirical strategy. Section 3.5 presents the main findings. In Section 3.6, I conclude with a discussion and policy implications.

[2]I do not run a two-stage specification. In this case, one needs to specify the functional form of the model in which efficiency is the dependent variable and mothers' education is the independent variable. One misspecification would lead to an incorrect efficiency measure (Ruggiero (1998)). Also, efficiency is a function of the expenditure per student, which is incorporated in the error term of the second stage. The estimated efficiency will be biased if the expenditure is correlated with mothers' education. This would be the case under the reasonable assumption that richer municipalities, where the percentage of mothers with a high school degree is higher, expend more per student.

3.2 Related Literature

The education production function is used to investigate the relationship between school inputs, such as per-pupil expenditure, and students' outputs, such as academic achievement (Pigott et al. (2012)). Coleman (1968) is the vanguard in that regard, setting the standards adopted since then. The author analyzes survey data from 600,000 students, 60,000 teachers, and 4,000 public schools in the United States, aiming to understand what are the determinants of students' capacity to learn. He then explores the influence of schools' amenities, students' family background and peers, per-pupil expenditure, and teachers' knowledge and practices in the classroom. The author concludes that family background is the main determinant of how well children learn, instead of schools' physical infrastructure or funding. His analysis is followed by a spate of studies that gather new data to explore the relationship between schools' inputs and outputs.

Hanushek (1981, 1986, 1996) analyzes hundreds of studies to understand the relationship between students' achievement and schools' expenditures, mostly determined by pupil-teacher ratios, and teachers' salaries (largely explained by their education and experience). The author concludes that additional dollars on traditional policies, such as reducing class sizes or hiring more qualified teachers, are unlikely to be matched by a significant increase in student achievement in the USA.

The weak association between increased per-pupil expenditure and students' performance raises questions on how efficiently public education expenditures are made, leading to extensive literature aimed to measure it. The most common methodologies employed for this task are Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). On the one hand, the main benefit of the DEA is that it is non-parametric, as there is no need to specify a functional form relating inputs to outputs. On the other hand, SFA, under the specification of a functional form, allows the differentiation of statistical noise from inefficiency (Ruggiero (2007)).

Schools, municipalities, or countries are the decision-making units (DMUs) in charge of the education process. In DEA and SFA, a production frontier is empirically estimated by comparing input-output combinations of several DMUs. The production units on the frontier are the ones that achieve the maximum output level given the inputs being used and the technology available. They provide a benchmark from which other DMUs in the sample are compared too, allowing the calculation of an inefficiency measure. In this sense, the drawn frontier depends on the sample chosen and does not necessarily mean that there is no room for improvement. Nonetheless, the more DMUs included in the sample, the closer the estimated frontier is to the true one (Mattos and Terra (2015a)).

Under a technical efficiency framework, the most common inputs are pupil-teacher ratio, school amenities, and school hours. Still, DMUs can be technically efficient but use inputs that might be expensive. In this context, a cost-efficiency analysis arises by adopting per-pupil expenditure as the main input (Afonso and Aubyn (2005)). Also, non-discretionary inputs, such as socioeconomic background, are commonly used as they play a crucial role in students' achievement, but cannot be changed, unless in the long run. On the output side, the most common ones are students' performance in reading and math, school dropout, grade promotion, and attendance rates (Hanushek (1981)).

Sutherland et al. (2007) employ a Data Envelopment Analysis to investigate the efficiency of the OECD countries with regard to 15-year-olds performance in reading, math, and sciences on the 2003 Programme for International Student Assessment (PISA). At the school level, the authors conduct a technical efficiency analysis in which the inputs are teaching staff and the number of computers per 100 students. According to their input-oriented estimates, schools could save up around 30% of their resources, and still achieve the same PISA if they move towards the efficient frontier. The output-oriented estimates indicate that schools could boost their PISA by up to 20% while holding per-pupil expenditure constant if they reach the frontier. Also, a cost-efficiency analysis is performed using country-level data with per-pupil expenditure as input. The authors estimate potential financial savings of almost 20%. Students' socioeconomic background is included as a non-discretionary input in all the analyses.

Afonso and Aubyn (2005) also employ a DEA on the OECD countries using the 2000 PISA results. Their study differs from Sutherland et al. (2007) as the inputs are total intended instruction time and pupil-teacher ratio. Their results are more conservative and show scope for savings of 14% of a boost in performance of 7%. Herrera and Pang (2005) use data for 140 countries from 1996 to 2002 to run a cost-efficiency DEA, in which the input is the aggregated public expenditures. The authors use several educational and health outcomes, such as primary and secondary enrollment, literacy rates, average years of schooling, graduation rates, learning scores, life expectancy at birth, and immunization rates. Their estimates show the potential to increase outputs by 10% to 30% if the countries use their resources efficiently.

The OECD countries are significantly different than Brazil. According to the Education at a Glance (2021), the USD 3,748 expenditure per primary education student in Brazil is approximately two and a half times lower than the USD 10,101 OECD average, a fact that is mainly explained by the gap in teachers' annual salaries (USD 25,366 versus USD 45,687). Even so, Menezes-Filho et al. (2009) also find evidence of a weak association between per-pupil expenditure in Brazil on 2005 students' performance in reading and math. Their study is followed by several efficiency analyses whose results suggest that Brazilian schools fail to use resources efficiently.

Rocha et al. (2013) assess the cost-efficiency of the Brazilian municipalities using a DEA in which the output is the Education Development Index (IDEB). They assume that municipalities have different technologies depending on their population size and, therefore, do not have the same education production function. The authors estimate five production frontiers using per-pupil expenditure and the education of students' mothers as the discretionary and non-discretionary inputs, respectively. The estimates show potential financial savings of 40% if municipalities use their resources efficiently.

Araújo Junior et al. (2019) estimate the technical efficiency of the Brazilian Northeastern municipalities and whether there was an improvement between 2007 and 2013. Similar to Rocha et al. (2013), the authors define five clusters of municipalities, based on population size, GDP per capita, quality of life, and economic development. They find evidence of an improvement in efficiency over the years. However, the authors conclude that more than half of the municipalities could boost their proficiency by at least 25% without increasing per-pupil expenditure if they move towards the efficient frontier. Also, the analysis indicates that smaller municipalities tend to be less efficient, similar to the results of Rocha et al. (2013) and Sousa et al. (2005).

Melo Castro et al. (2017) employ a stochastic frontier approach to estimate an education production function for Brazilian municipalities. The output is the percentage of students with an adequate level of proficiency in reading and math. The inputs are school daily hours, percentage of teachers with undergrad, students' household infrastructure and parents' support, municipalities' illiterate rates, inequality levels, and urbanization rates. The authors estimate the prices of the labor force, capital, and other administrative expenses based on the total expenditure declared by the municipality in each of these categories and what these amounts can afford in terms of schools' staff, and infrastructure. The estimates show a cost-inefficiency of almost 15% and that higher efficiency levels could be achieved if schools increase the length of the school day and at the same time increase the pupil-teacher ratio.

3.3 Data

In Brazil, the 26 states, the Federal district, and the 5,570 municipalities share the responsibilities for the provision of public education. According to the 1988 Constitution, the municipal governments should give priority to early childhood and primary (grades 1 to 5) and lower secondary education (grades 6 to 9), and the state authorities to primary, lower, and upper secondary education (grades 10 to 12). The municipalities can then have schools managed by the municipal government or/and schools managed by the state government. The first group of schools is the locally-managed ones, and the second group of schools is the state-managed ones.

Article 211 - 1988 Brazilian Constitution
§ 2o Os Municípios atuarão prioritariamente no ensino fundamental e na educação infantil.
Municipalities will give priority to providing early childhood, primary, and lower secondary education.
§ 3o Os Estados e o Distrito Federal atuarão prioritariamente no ensino fundamental e médio.
States will give priority to providing primary, lower, and upper secondary education.

According to the 2019 Census of Education, 5,555 municipalities out of 5,570 Brazilian municipalities have at least one primary education school managed by the respective local authority. 2,154 municipalities out of 5,570 have primary schools managed by the state governments. Under this organization, municipal governments have almost 85% of public primary education students enrolled in schools under their management.

Since 1995, all private and public schools offering primary and secondary education participate in the annual Census of Education. The Census is implemented by the National Institute of Educational Studies and Research (INEP), a research agency under the Brazilian Ministry of Education.[3] The Census collects information on (i) school facilities, such as libraries, sports courts, and science and computer labs; (ii) school infrastructure, such as filtered water, electricity, and internet access; (iii) social services, for example, school transportation and provision of meals; (iv) students, such as their sex, color of the skin, age, physical disabilities or mental illness, grade level, instruction time per day, class-size, subjects they are enrolled in, grade promotion, repetition and dropout; and (v) teachers, such as their educational attainment, age, physical disabilities, subjects taught, and classes they are in charge of. The agency then analyses the Census to disclose educational indicators at school, municipality, state, and national levels. The indicators are disclosed by the type of school: private, local, state, and federal-managed schools. Therefore, I observe for each municipality, for example, the average dropout rate of locally-managed schools, and of state-managed ones.

Every two years, INEP applies a proficiency assessment, *Prova Brasil*, to fifth graders of public and private schools, a test that is within the scope of the Education Assessment System (SAEB).[4] For fifth-graders, the exam has a scale ranging from 0 to 325, and a math scale ranging from 0 to 350 (SAEB scale). Children also answer a socioeconomic questionnaire with information on their household infrastructure; parents' educational attainment; incentives from their family to pursue an education; time watching TV, on the internet, reading books, and doing homework; if they already dropped out or repeated a grade; and if they did kindergarten. The socioeconomic information can also be aggregated at the municipality level by schools' level of administration, for example, the percentage of fifth-graders in locally-managed schools whose mothers finished high school.

Data from the Census of Education and *Prova Brasil* are used to calculate the National Education Development Index (IDEB), the most important educational indicator in Brazil. IDEB monitors students' grade promotion and proficiency in reading and math.[5] To compute the index, the students' Portuguese and math performance is computed on a scale from 0 to

[3]INEP stands for *Instituto Nacional de Estudos e Pesquisas Educacionais.*

[4]Schools with at least 20 students enrolled. Proficiency tests are also applied to students in the ninth grade and in the last grade of high school. SAEB stands for *Sistema de Avaliação da Educação Básica.*

[5]IDEB stands for *Índice de Desenvolvimento da Educação Básica.*

10. This score is then multiplied by the students' grade promotion rates (on a scale from 0 to 1) to obtain the IDEB at school, municipal, state, and national levels. Therefore, the index ranges from 0 to 10. For primary education, the index is the product of the standardized performance of fifth-graders and the average grade promotion from first to fifth grades. Central and local governments use IDEB to monitor the improvement in the quality of public and private education in Brazil and compare schools' performance within and between municipalities, and also between private, local, and state-managed networks. For example, at the municipal level, I observe the IDEB of locally-managed schools, and of state-managed ones.

The Information System on Expenditures in Education (SIOPE) is released by the National Fund for Education Development (FNDE), an agency under the Brazilian Ministry of Education.[6] Municipal and state governments have to report to SIOPE their total expenditures in education, such as teachers' salaries, school infrastructure, textbooks, and an estimate of per-pupil expenditure. The information needs to be reported by the level of education: i) preschool and kindergarten, ii) primary and lower secondary education (grades 1 to 9), and iii) upper secondary education (grades 10 to 12). SIOPE then discloses the data at the state level for the 26 state governments and the Federal District, information that shows the average per-pupil expenditure in state-managed schools. SIOPE also disclosed the data at the municipal level for the 5,570 local authorities, information that shows per-pupil expenditures in locally-managed schools. Hence, it is possible to observe per-pupil expenditure across all state and municipal governments.

To estimate the efficiency of public primary education expenditures in Brazil, I restrict the sample to schools managed by the municipal governments, as almost all municipalities provide first to fifth grades (5,555 out of 5,570) and are in charge of approximately 85% of enrollment in this level of education. The outcome is 2019 IDEB, which I consider that it is a function of the per-pupil expenditure and students' socioeconomic background. As the performance of fifth graders is influenced by the investments made since they joined the school, I work with the per-pupil expenditure in 2019, when students were in fifth grade, in 2018 (fourth grade), in 2017 (third grade), in 2016 (second grade), and 2015 (first grade).[7] To account for students'

[6] SIOPE stands for *Sistema de Informações sobre Orçamentos Públicos em Educação*. FNDE stands for *Fundo Nacional de Desenvolvimento da Educação Básica*.

[7] This assumption is valid for students that did not repeat any grade, which is plausible for this level of

background once it plays an important role in their performance, I use as a non-discretionary input the percentage of students whose mothers finished high school (Table 3.1).

Table 3.1: DMUs, output, and inputs of the DEA

DMUs
Municipal governments of 5,570 Brazilian municipalities
Output
2019 IDEB at the municipal level, which is an average of the IDEB of the schools managed by the municipal government.
Inputs
Discretionary
Expenditure per student made by the municipal governments on their locally-managed schools (2015-2019)
Non-discretionary
Percentage of students whose mothers finished high school (2019)

Brazil is a continental country where the cost of living significantly changes from north to south and in metropolitan, urban and rural areas. To account for that the same amount of per-pupil expenditure does not mean the same depending on the region, I adjust this variable by an estimate of the cost of living across the whole country (Table C.1). If the municipality belongs to a metropolitan area, I multiply the per-pupil expenditure by the respective cost of living. For the remaining municipalities, I first calculate the percentage of enrollments in urban and rural areas and use them as a weight when multiplying the per-pupil expenditure by the respective cost of living in these areas.[8]

According to 2019 information disclosed by the Brazilian Institute of Geography and Research (IBGE), municipalities have an average population size of 37 thousand. Around 25% of them had up to 5 thousand inhabitants, and only 1% has more than 400 thousand.[9] On the one hand, more populated municipalities might face more challenges to manage their school systems, which are likely to be more complex as they have more schools under their management. On the other hand, some small municipalities might have less qualified personnel in charge of their educational policies and are more likely to be captured by local elites. To account for the fact that these

education once the repetition rate is low (0.6% in 2019 for locally-managed schools).

[8] After that, I exclude from the analysis the municipalities where the expenditure per student is in the first or in the 99th percentile of the expenditure distribution.

[9] IBGE stands for *Instituto Brasileiro de Geografia e Estatística*.

governments have different education production functions, I work with five groups, following the IBGE definition when dividing municipalities into clusters according to their population size.

1. Municipalities that have up to 5 thousand inhabitants (22.3%).
2. Municipalities that have between five thousand and fifty thousand (65.6%).
3. Municipalities that have between fifth thousand and 100 thousand (6.3%).
4. Municipalities that have between 100 thousand and 500 thousand (5%).
5. Municipalities with more than 500 thousand (0.8%).

The descriptive statistics for per-pupil expenditure and students' performance and socioeconomic background are shown in Table C.2. To perform the linear programming of the Data Envelopment Analysis, I need to work on municipalities in which all these variables are available. This reduces the sample size from 5,555 to 4,807 municipalities (Table C.3). Therefore, my model includes 86% of the Brazilian municipalities.

3.4 Empirical Strategy

The Data Envelopment Analysis (DEA) is a non-parametric methodology for efficiency analysis first introduced by Charnes et al. (1978). The idea is to assess the relative efficiency of a set of Decision Making Units (DMUs), which are the production units under evaluation, such as schools, hospitals, municipalities, or countries. A linear programming model is employed to compare the inputs used by the DMUs, such as the per-pupil expenditure and doctors per capita, with the outputs achieved, such as students' proficiency and children's mortality rates.

DEA framework can easily accommodate multiple outputs and inputs and it is not necessary to specify any functional form (Sutherland et al. (2007)). Also, the methodology assumes that linear combinations of observed input-output bundles are feasible (Herrera and Pang (2005)). Considering the technology available, the output-oriented efficiency measures the difference between what the DMU achieves and what would be possible to achieve if it operates on the production frontier, that is, the shortfall of outputs for a given level of inputs. The input-oriented efficiency measures the difference between the inputs being used and the minimum amount that would be enough to reach the same output level if the DMU operates on the frontier, that is, the excess of input consumption (Mattos and Terra (2015a)).

DEA specifications can have constant returns to scale (CRS), which assumes proportionality between outputs and inputs, or variable returns to scale (VRS), a relaxation of the previous assumption as it allows that the increase in the inputs does not need to be followed by a proportional increase in outcomes for the DMU to be considered efficient. The VRS is more adequate for this analysis since, given a specific increase in one of the inputs, it is more feasible for a low-performer DMU to increase its education outcomes compared to a unit that already has a high performance.

Consider a VRS DEA with n DMUs in the reference set, i inputs, and s outputs. An input-oriented approach consists of solving, for each decision-making unit k under evaluation, the following problem:

$$\text{Min}\,\theta_k, \text{ subject to:}$$

$$\sum_{j=1}^{n} \alpha_j x_{ij} \leq \theta_k x_{ik}, \; \forall\; i \in 1, 2, .., m$$

$$\sum_{j=1}^{n} \alpha_j y_{rj} \geq y_{rk}, \;\; \forall\; r \in 1, 2, .., s$$

$$\sum_{j=1}^{n} \alpha_j = 1, \; \alpha_j > 0$$

In which j is the subscript for all the DMUs in the sample and assumes values from 1 to n. x_{ij} and x_{ik} represent the input i being used by the DMU j and k, respectively; and y_{rj} and y_{rk} represent the output r reached by DMU j and k, respectively. Intuitively, the model performs n minimization problems, one for each DMU, with $m+s+1$ restrictions to check whether there is a linear combination of inputs in the reference set, weighted by α_j, that is inferior to the one being used by the DMU k and that still produces more or the same amount of outputs. θ_k represents the share of inputs being used efficiently by DMU k, the input-oriented efficiency score.

The DMU k is considered efficient if there is no combination of inputs in the reference set of DMUs that uses fewer inputs and still produces more outputs than it. In this case, θ_k assumes a value of 1. Therefore, it is possible to observe that, by construction, at least one DMU in the sample is classified as efficient and provides a benchmark to which others are compared. For inefficient DMUs, there is a linear combination of inputs in the reference set that is inferior to the one being used by it but that produces the same or more outputs. In this case, $\theta_k < 1$, and $1 - \theta_k$ represent the proportion of inputs that should be reduced for the DMU k to become efficient. This model has a strong assumption as it assumes that all the inputs can decrease. For a more accurate analysis, one can specify non-discretionary inputs, the ones that, at least in the short run, cannot be changed.

The equation below shows the minimization problem under an output-oriented approach. In this case, if there is not a linear combination of DMUs in the reference set that employs fewer inputs than unit k and produces more outputs, θ_k assumes value 1. Otherwise, $\theta_k < 1$, and $\frac{1}{\theta_k}$ is the proportional expansion of outputs needed for the DMU k reach the frontier. Hence, θ_k is the output-oriented efficiency score.

$$
\begin{aligned}
&\text{Min}\, \theta_k, \text{ subject to:} \\
&\sum_{j=1}^{n} \alpha_j x_{ij} \leq x_{ik}, \; \forall\ i \in 1, 2, .., m \\
&\sum_{j=1}^{n} \alpha_j y_{rj} \geq \frac{y_{rk}}{\theta_k}, \;\; \forall\ r \in 1, 2, .., s \\
&\sum_{j=1}^{n} \alpha_j = 1, \; \alpha_j > 0
\end{aligned}
$$

The DEA constructs a production frontier containing all the DMUs for which $\theta = 1$. They provide a benchmark measure for each inefficient unit. The model also provides a linear combination of inputs that represents the efficient amount that the DMU k under evaluation needs to employ to emulate the ones on the frontier and achieve $\theta = 1$.

However, this approach has one caveat when it comes to the presence of outliers in the sample. Municipalities with very low expenditure per student would make us overestimate the inefficiency of the other DMUs in the reference set. To deal with this, I run a super-efficiency (SE) model to check whether the DMUs that achieved $\theta = 1$ in the first minimization problem are actually outliers instead of good examples that other production units could emulate.

Under this approach, each one of these DMUs is excluded from the sample, one at a time, and a new minimization is performed with the remaining observations. For an efficient DMU k identified in the first minimization problem, the input-oriented VRS-SE model consists in solving:

$$
\begin{aligned}
&\text{Min}\,\theta_k, \text{ subject to:}\\
&\sum_{\substack{j=1\\j\neq k}}^{n} \alpha_j x_{ij} \leq \theta_k x_{ik},\ \forall\ i \in 1, 2, .., m\\
&\sum_{\substack{j=1\\j\neq k}}^{n} \alpha_j y_{rj} \geq y_{rk},\ \ \forall\ r \in 1, 2, .., s\\
&\sum_{\substack{j=1\\j\neq k}}^{n} \alpha_j = 1,\ \alpha_j > 0
\end{aligned}
$$

For the efficient DMUs in the first minimization problem, θ_k under the super-efficiency specification will assume a value higher than 1. If the θ_k is equal to 1.5, this indicates that the DMU k could increase its inputs up to 50% and would still be considered efficient when compared to the DMUs in the reference set. The higher the θ_k, the more evidence that the DMU k is an outlier. One can use different thresholds to select the outliers. In this chapter, I present the results for 1.3, 1.2, and 1.1. Therefore, these models exclude DMUs that could increase input levels by 30%, 20%, and 10%, respectively, and they would still be efficient when compared to the remaining units in the reference set. Banker et al. (2017) finds evidence that supports the use of this methodology to identify outliers. Also, super-efficiency has been employed by several authors such as Banker and Gifford (1988), Andersen and Petersen (1993), Tone (2002), Banker and Chang (2006), and Avkiran (2011). One caveat is that the choice of the thresholds, as documented in the literature, is subjective (Ahamed et al. (2015)).

Another very important assumption of DEA is that the DMUs in the reference set are similar to each other, share the same production frontier, and have comparable technologies. These are strong assumptions given that the production units might have diverse infrastructure and physical and human capital (Boueri (2015)). To overcome this threat, I assume that the DMUs operate under different production functions by dividing them into groups that are somehow comparable to each other and running a DEA for each one of them. I define five clusters of municipalities based on population size as described in Section 3.3.

First, for each one of the five groups, I run a DEA model that measures the efficiency of the DMU k relative to the common frontier of its group, the *Group-Frontier*. The comparison of the efficiency scores of the five clusters is not valid as it is assumed that the groups have different production environments. Second, in a hypothetical scenario in which all the municipalities have the same technology, I run a DEA for the five groups together, the *Meta-Frontier*. The distance between the group and the meta-frontiers is an estimated measure of the technological gap existent between the DMUs (Wongchai et al. (2012)). In this sense, the DMUs in the group frontier would move to the meta frontier if they could increase the productivity of the inputs being used (Mattos and Terra (2015b)).

To assess the efficiency of public primary education expenditures in Brazil, the 2019 primary education IDEB is the only output of the model. The index represents the main variables used in the educational production process as it combines students' proficiency with grade promotion, (Hanushek (1996)). As Hanushek (1986) documents, past investments made on children influence their current level of achievement, I work with five discretionary inputs: the per-pupil expenditure made in 2019, 2018, 2017, 2016, and 2015, all expressed in 2020 BRL. It is also well documented that students' background significantly affects their performance. In this sense, municipalities can have the same per-pupil expenditure, but different performance, and a simple comparison would not be adequate, as they might have to deal with distinct socioeconomic contexts. It is notably easier for a municipality whose students' mothers have higher education achievement to reach higher reading and math proficiency for a given level of expenditure. For that reason, I use as a non-discretionary input the percentage of students whose mothers finished high school (Coleman (1968), Hanushek (2006)).

The incorporation of a non-discretionary input in the model is more adequate than running a two-stage model in which the first step is to estimate the efficiency of the DMUs based on the discretionary inputs, and the second step is to run the estimated efficiency against mothers' education. This last approach requires a functional form specification for the second stage of regression. In this case, misspecification leads to an inaccurate measurement of efficiency (Ruggiero (1998)). Also, mothers' education might be correlated with per-pupil expenditure, as more affluent municipalities (those with children with better socioeconomic backgrounds) might spend more, leading to biased estimates for inefficiency.

3.5 Results

3.5.1 Input-oriented

Table 3.2 shows the results of the input-oriented approach. The number of DMUs that reach the efficient frontier increases in the super-efficiency models as these specifications exclude from the analysis municipalities with low levels of per-pupil expenditure, thee ones that are more likely to be outliers instead of examples of good practices that other municipalities could emulate. Therefore, more DMUs in the reference set reach the frontier. Table C.3 shows that these specifications do not exclude more than 2% of the sample, suggesting that only a few municipalities are classified as outliers. Not considering these identified outliers in the analysis increases the number of benchmark units up to 17% (from 267 to 313 municipalities). I show the results for three thresholds: 1.3, 1.2, and 1.1 as explained in Section 3.4. The lower the threshold, the less likely I underestimate the efficiency of the municipalities. Therefore, to be conservative, the most adequate specification is shown in Column IV.[10]

My results suggest that there are 313 municipalities on the group-frontier ($\theta = 1$), which represents 7% of the DMUs in the sample. These municipalities have an efficient combination of per-pupil expenditure achieving the maximum possible performance given their current level of expenditures compared to other local authorities in their group sample. 50% are in the north and northeast of Brazil (Table C.4). The results show that a very low percentage of Brazilian municipalities are on the efficient frontier, suggesting significant space for fiscal savings.

The efficient scores indicate that, on average, municipalities efficiently use between 72% and 83% of public primary education resources, depending on their size.[11] The results suggest that if the local authorities could emulate the educational policies adopted by the frontier

[10] Since the inefficiency of the municipality is given by 1-θ, in which θ is the efficiency estimate of the model, if I underestimate the efficiency, I overestimate the inefficiency. That is why, to be conservative, I opt to work with the specification that has a threshold of 1.1.

[11] Table Table C.5 shows the input-oriented efficiency scores in 2013. The results suggest only a slight increase in efficiency over time. Also, Figure C.2 shows that the higher the efficiency in 2013, the higher the efficiency in 2019.

municipalities, the ones that are somehow similar to them, there would be a fiscal space of at least 86 billion BRL. This amount is more than 16 times the approved budget for the National Fund for Education Development (FNDE) in 2022. FNDE is a federal government fund that finances several programs targeting primary, lower and upper secondary education in Brazil. The fund is in charge of children's textbooks, free meals in schools, school transport, and school infrastructure and repairs, among others.

To give an example of what that amount represents in terms of educational policies, an estimated 450 million BRL would be needed to finance an initiative to reduce the dropout that could target all the 28.6 million students in public primary, lower, and upper secondary education in Brazil. An early warning system strategy aims to reduce dropouts by identifying high-risk students based on several factors that affect their decision to remain in school, such as proficiency levels, absence, teenage pregnancy, and whether they work or not. This type of intervention was implemented in Guatemala in 2017 and Haimovich et al. (2021) estimates a decrease in dropout of 9%.[12]

One may wonder whether the sample of municipalities only includes low-performers, which would make me underestimate the inefficiency in Brazil. I do not find evidence that this is the case. The frontier has municipalities that are nationally recognized for their good educational outcomes. For example, the state of Ceará has 19 municipalities on the frontier, one of them is Sobral. This municipality is well-known for its educational practices and has educational outcomes, comparable to what is achieved by OECD countries. Also, 10% of the municipalities in the sample have IDEB, the output under evaluation in the model, 20% higher than the target established by the Ministry of Education. Therefore, the sample includes several DMUs that could be examples of good practices, suggesting that I do not underestimate the inefficiency.

Except for the municipalities that have between 5 and 50 thousand inhabitants, all the other clusters have group-frontier scores significantly higher than the meta-frontier ones. This indicates that municipalities with up to 5 thousand inhabitants and those with more than 50 thousand have an educational production function assumed to have a technological gap when compared with municipalities that have between 5 and 50 thousand inhabitants, the ones that mostly made up the meta-frontier. On the one hand, bigger municipalities might face more

[12] The program had an estimated cost of 3 USD per student. I use the exchange rate of 5.22 BRL (exchange rate as of September 6, 2022).

challenges to manage their school systems, which are likely to be more complex as they have more schools under their management. On the other hand, some small municipalities might have less qualified personnel in charge of their educational policies and are more likely to be captured by local elites.

Table 3.2: Input-oriented efficiency scores (2019)

	Group Frontier				Meta Frontier			
	Baseline	Robustness				Robustness		
	Without exclusions	Excluding outliers Thresholds			Without exclusions	Excluding outliers Thresholds		
		1.3	1.2	1.1		1.3	1.2	1.1
	I	II	III	IV	I	II	III	IV
Number of benchmarks								
Up to 5k	54	69	67	71	14	10	7	6
Between 5k and 50k	119	121	132	137	111	113	124	129
Between 50k and 100k	50	51	51	59	9	8	9	3
Between 100k and 500 k	32	35	35	36	12	10	10	5
More than 500k	12	10	10	10	1	0	0	0
Total, Brazil	267	286	295	313	147	141	150	143
Input-oriented efficiency scores								
Up to 5k	0.69	0.71	0.71	0.72	0.59	0.59	0.59	0.60
Between 5k and 50k	0.73	0.73	0.73	0.74	0.72	0.73	0.73	0.74
Between 50k and 100k	0.81	0.81	0.81	0.83	0.73	0.73	0.73	0.74
Between 100k and 500 k	0.78	0.79	0.79	0.80	0.71	0.71	0.71	0.72
More than 500k	0.78	0.81	0.81	0.83	0.60	0.57	0.58	0.56
Total, Brazil					0.70	0.70	0.70	0.71
Potential fiscal space, in billion 2020 BRL								
Up to 5k	3.37	3.15	3.10	2.97	4.41	4.39	4.33	4.23
Between 5k and 50k	33.80	33.72	33.35	32.50	34.20	34.06	33.71	32.71
Between 50k and 100k	9.29	9.03	8.98	7.72	12.78	12.70	12.60	12.09
Between 100k and 500 k	23.29	22.15	22.15	20.44	29.86	29.65	29.57	28.28
More than 500k	30.78	25.97	25.97	22.81	47.33	46.78	46.62	45.40
Total, Brazil	100.54	94.01	93.55	86.45	128.58	127.58	126.84	122.72

Source: The authors' estimate is based on 2019 *Prova Brasil*/INEP, and 2015-2019 SIOPE. Municipal governments. The columns without exclusions show the number of benchmark municipalities before the exclusion of outliers. The super-efficiency columns indicate the number of benchmark DMUs in the models that exclude outliers. I show the results for three thresholds: 1.3, 1.2, and 1.1. In the first threshold, the model excludes from the analysis the municipalities whose inputs could increase up to 30% and they would still be on the frontier. These percentages are 20% and 10% for the remaining specifications. The group frontier does not have efficient scores at the national level because the frontiers of each group of municipalities are not comparable. See Section 3.4.

An interesting result of DEA is being able to compare the municipalities used as benchmarks with the ones that do not reach the production frontier. One DMU can have several benchmark municipalities and the model attributes a weight for each one of them. Table 3.3 shows examples of municipalities that are not on the frontier. For example, the benchmark DMUs for the municipality of Cajuri, in the state of Minas Gerais, are Arantina, Augusto de Lima, and Senador Cortes, all in the same state and with less than 5 thousand inhabitants. I can use the weights shown in Table 3.3 and the level of expenditure per student in these benchmark municipalities to estimate the efficient amount of expenditure per student in Cajuri. Their accumulated expenditure per primary education student between 2015 and 2019 was 36.5 thousand BRL. According to DEA estimates, if the municipality reaches the frontier, it would be able to achieve the same educational outcome by spending 30 thousand BRL. Identifying the benchmark municipalities could contribute to the dialogue between municipal governments to share policy interventions aimed at increasing students' proficiency.

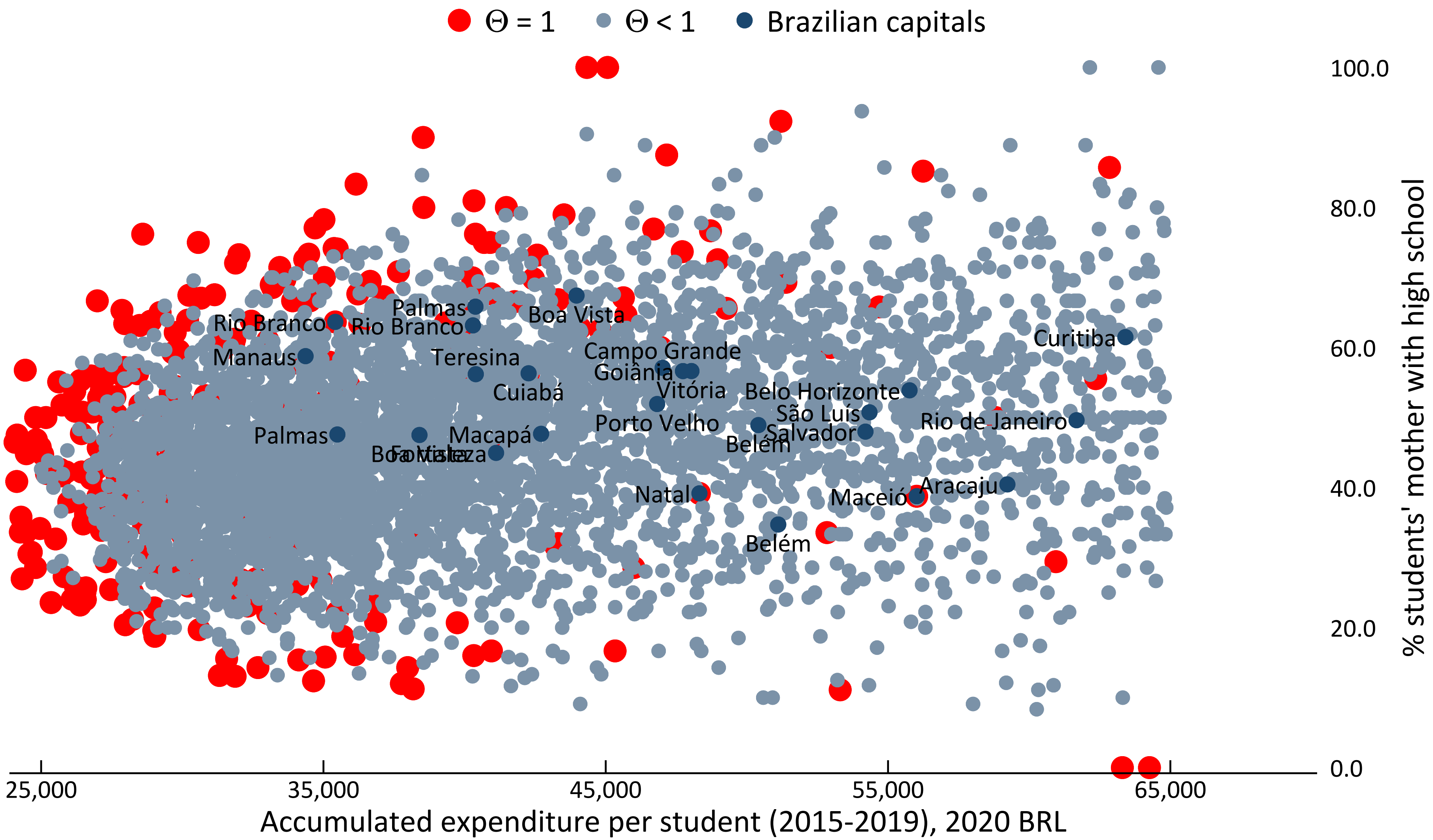
Table 3.3: Example of benchmark DMUs (2019)

	Input-oriented efficiency score	Benchmarks/weights		
Lorena/SP	0.76	Pedro Leopoldo/MG (43.2%)	Viana/ES (45.5%)	São José do Rio Pardo/SP (11.2%)
Caxias/MA	0.92	Itacoatiara/AM (49.6%)	Sabará/MG (50.3%)	0.00
Emas/PB	0.79	São Pedro dos Crentes/MA (8.59%)	Jurema/PI (40.8%)	Augusto de Lima/MG (50.5%)
Tibau/RN	0.49	Arantina/MG (55.5%)	Augusto de Lima/MG (35.0%)	Senador Cortes/MG (9.34%)
Escada/PE	0.94	Buriticupu/MA (.712%)	Palmares/PE (19.8%)	Timbaúba/PE (79.4%)
Cajuri/MG	0.85	Arantina/MG (45.6%)	Augusto de Lima/MG (14.6%)	Senador Cortes/MG (39.7%)
Acará/PA	0.96	Monte Alegre/PA (41.8%)	Pedro Leopoldo/MG (39.4%)	Viana/ES (18.6%)
Pinhal/RS	0.74	Baraúna/PB (73.5%)	Congo/PB (9.83%)	Olímpio Noronha/MG (16.6%)
Tupã/SP	0.69	Coruripe/AL (2.32%)	Pedro Leopoldo/MG (55.5%)	Andradina/SP (42.1%)
Cumbe/SE	0.89	Vila Flor/RN (19.1%)	Assunção/PB (53.4%)	Joaquim Felício/MG (27.4%)
Maraã/AM	0.82	Vale do Anari/RO (20.5%)	Uarini/AM (70.2%)	Barra de Guabiraba/PE (9.16%)
Lins/SP	0.59	Coruripe/AL (18.1%)	Pedro Leopoldo/MG (68.1%)	Andradina/SP (13.6%)
Calmon/SC	0.83	São Pedro dos Crentes/MA (9.29%)	Assunção/PB (49.4%)	Congo/PB (41.2%)
Maruim/SE	0.65	Agricolândia/PI (67.0%)	Bonito de Santa Fé/PB (4.76%)	Tanguá/RJ (28.1%)
Franca/SP	0.70	Sobral/CE (48.0%)	Linhares/ES (51.6%)	Taboão da Serra/SP (.309%)
Jataí/GO	0.57	Sobral/CE (.681%)	Santa Luzia/MG (14.9%)	Francisco Morato/SP (84.3%)

Source: The authors' estimate is based on 2019 *Prova Brasil*/INEP, and 2015-2019 SIOPE. Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1. Input-oriented VRS.

Figure 3.1 shows the distribution of the accumulated expenditure per primary education student, from their first to fifth grade, and the percentage of students whose mothers finish high school. As expected, the municipalities that reach the frontier ($\theta_k = 1$) are the ones that, given a more favorable family background, are able to achieve high performance even with low levels of expenditure. Also, municipalities, where students have a less favorable background, can also be efficient even though they have higher per-pupil expenditure.

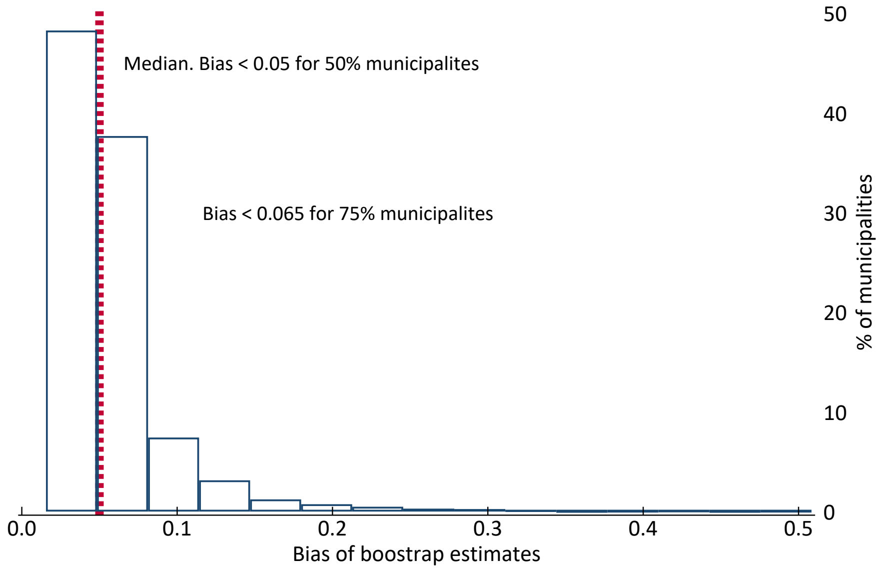
Figure 3.1: Distribution of per-pupil expenditure and socioeconomic background (2019)



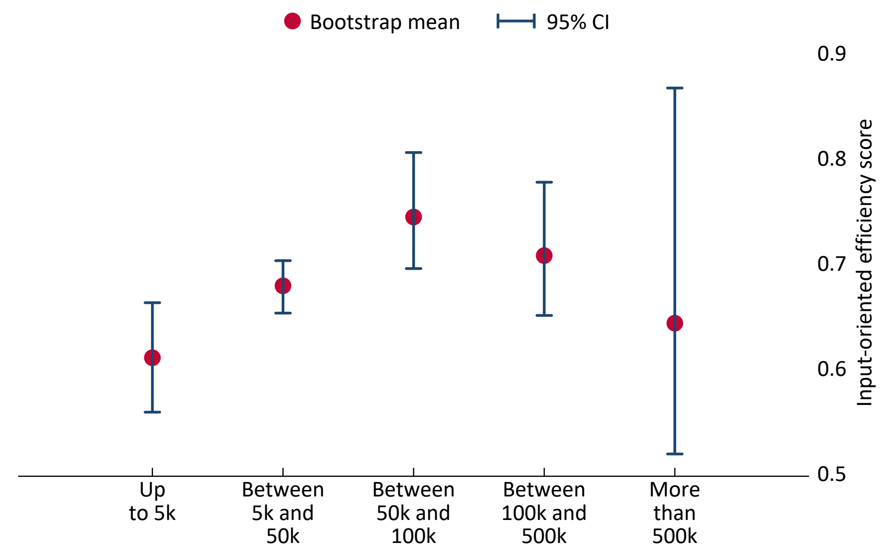
Note: The authors' estimate is based on 2019 *Prova Brasil*/INEP, and 2015-2019 SIOPE. Input-oriented VRS. Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1. The accumulated per-pupil expenditure is the sum of per-pupil expenditure from 2015 to 2019.

To further enrich the analysis, I run bootstrap models to calculate the confidence intervals for the input-oriented efficiency scores following Simar and Wilson (1998). The authors suggest this methodology as it is an easier way to analyze the sensitivity of efficiency scores relative to the sampling variations of the estimated frontier. On the one hand, panel (a) in Figure 3.2 shows that for 50% of the municipalities the difference between the input-oriented efficiency score of the original sample and the average efficiency score of bootstrapped samples is lower than 0.05. For 75% of the municipalities, the bias is lower than 0.065. These results indicate that overall the input-oriented efficiency scores are robust to sampling variations. On the other hand, panel (b) in Figure 3.2 shows that the estimates for municipalities with more than 500 thousand inhabitants are sensitive to sampling variations, which could be due to the low sample size (only 38 municipalities as shown in Table C.3).

Figure 3.2: Boostrap bias and confidence intervals (2019)



(a) Boostrap bias



(b) Boostrap 95% CI

Note: Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1. 2000 replications. I run five models, one for each cluster.

Section 3.3 shows that municipal governments are in charge of 85% of the public primary education enrollments. I run one additional specification to assess whether the efficiency scores of these DMUs are sensitive to the inclusion of the state governments in the analysis. In this case, in addition to the 4,807 municipalities, I include 27 DMUs in the analysis, each one representing one Brazilian state. The inputs of these DMUs are the per-pupil expenditure made by the state governments in the schools under their management. The output is the average IDEB of the students enrolled in these schools.

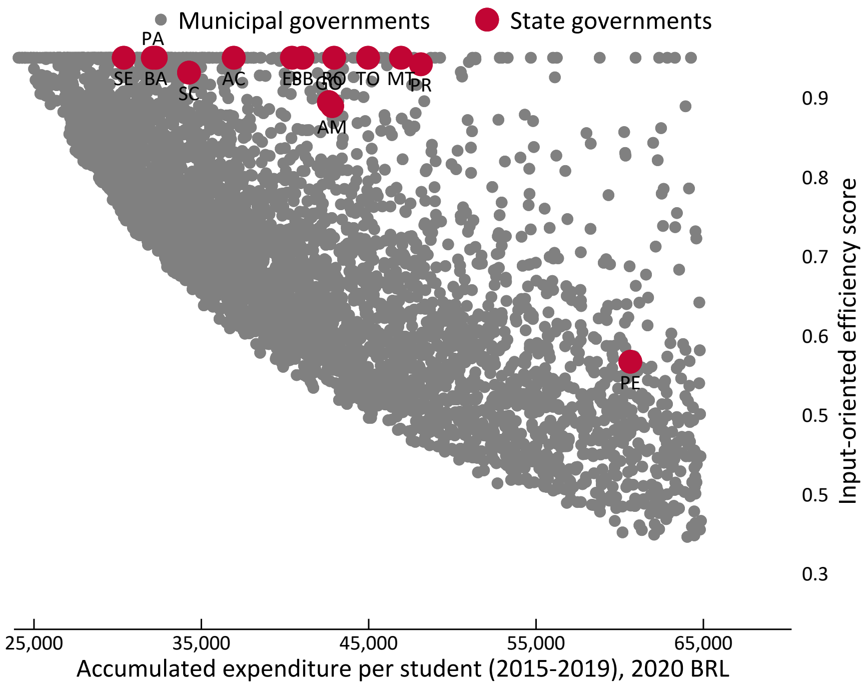
Figure 3.3 shows that the distribution of the estimated efficiency for the municipal governments does not change with the inclusion of the state-managed governments in the analysis. In panel (a), I show the distribution of expenditure per student versus the group-frontier input-oriented efficiency score. In this case, since all 27 states have more than 500 thousand inhabitants, their group-frontier analysis only includes municipalities that also have more than 500 thousand inhabitants. Their group-frontier efficiency score has an average of 88%, which is higher than the estimated efficiency for municipalities that are in the same group (72%). The results suggest that state governments tend to be more efficient than municipalities with more than 500 thousand inhabitants.

However, when compared to small municipalities, state governments have low-efficiency scores, except for Ceará. Figure 3.3 shows in panel (b) the distribution of meta-frontier efficiency scores. In this case, as explained in Section 3.4, all DMUs are assumed to have the same production function. This result is in line with the ones presented in Table 3.2, in which I observe that more populous DMUs have lower meta-frontier efficiency scores. State governments, similar to bigger municipalities, face more challenges to manage their school systems.

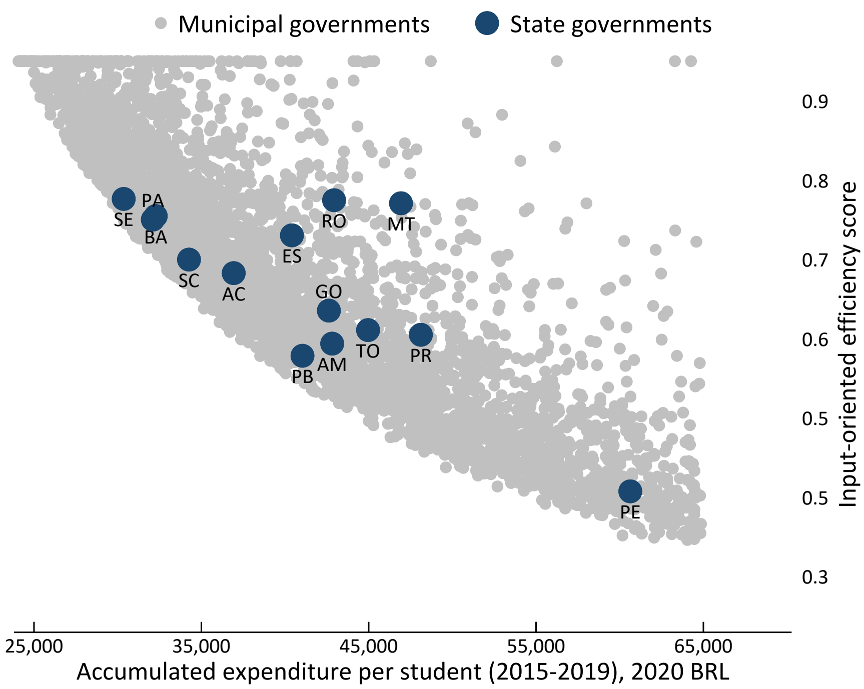
3.5.2 Output-oriented

Table 3.4 shows the results of the output-oriented approach. Given the current per-pupil expenditure, this model estimates what would be the potential performance of the students if municipalities were more efficient in the allocation of public resources. The estimates indicate that if the municipalities operated in the frontier, the national primary education IDEB could potentially reach 8, an increase of more than 40% when compared to 5.7 achieved by the students in 2019.

Figure 3.3: Distribution of input-oriented efficiency scores, municipal and state governments (2019)



(a) Group-Frontier



(b) Meta-Frontier

Note: Municipal and state governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1.

Table 3.4: Output-oriented efficiency scores (2019)

	Group Frontier				**Meta Frontier**			
	Baseline	**Robustness**				**Robustness**		
	Without exclusions	**Excluding outliers Thresholds**			**Without exclusions**	**Excluding outliers Thresholds**		
		1.3	**1.2**	**1.1**		**1.3**	**1.2**	**1.1**
	I	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
Number of benchmarks								
Up to 5k	54	69	67	71	14	10	7	6
Between 5k and 50k	119	121	132	137	111	113	124	129
Between 50k and 100k	50	51	51	59	9	8	9	3
Between 100k and 500 k	32	35	35	36	12	10	10	5
More than 500k	12	10	10	10	1	0	0	0
Total, Brazil	267	286	295	313	147	141	150	143
Output-oriented efficiency scores								
Up to 5k	0.77	0.79	0.80	0.80	0.68	0.68	0.68	0.68
Between 5k and 50k	0.68	0.68	0.69	0.69	0.68	0.68	0.69	0.69
Between 50k and 100k	0.79	0.80	0.80	0.81	0.69	0.69	0.69	0.69
Between 100k and 500 k	0.83	0.83	0.83	0.83	0.71	0.71	0.71	0.71
More than 500k	0.89	0.90	0.90	0.90	0.67	0.66	0.66	0.65
Total, Brazil					0.68	0.68	0.69	0.69
Potential IDEB								
Up to 5k	7.72	7.52	7.47	7.42	8.71	8.72	8.72	8.72
Between 5k and 50k	8.31	8.32	8.28	8.22	8.36	8.35	8.31	8.26
Between 50k and 100k	7.36	7.35	7.34	7.26	8.39	8.39	8.35	8.38
Between 100k and 500 k	7.28	7.24	7.24	7.26	8.45	8.45	8.43	8.44
More than 500k	6.44	6.43	6.43	6.46	8.71	8.84	8.82	8.90
Total, Brazil	8.06	8.02	7.99	7.93	8.44	8.43	8.41	8.37

Source: The authors' estimate is based on 2019 *Prova Brasil*/INEP, and 2015-2019 SIOPE. Municipal governments. The columns without exclusions show the number of benchmark municipalities before the exclusion of outliers. The super-efficiency columns indicate the number of benchmark DMUs in the models that exclude outliers. I show the results for three thresholds: 1.3, 1.2, and 1.1. In the first threshold, the model excludes from the analysis the municipalities whose inputs could increase up to 30% and they would still be on the frontier. These percentages are 20% and 10% for the remaining specifications. The group frontier does not have efficient scores at the national level because the frontiers of each group of municipalities are not comparable. See Section 3.4.

3.6 Discussion and Conclusion

In this chapter, I employ a Data Envelopment Analysis to investigate the efficiency of public primary education expenditures in Brazil. Under the assumption that the municipalities have distinct education production frontiers depending on their populational size, the group-frontier results indicate that the municipal governments efficiently use between 72% and 83% of their educational resources. This result suggests that if these local authorities could reach the production frontier, there would be a fiscal space of at least 86 billion BRL. This amount is more than 16 times the approved budget for the National Fund for Education Development (FNDE), the most important federal fund that finances several programs targeting primary, lower, and upper secondary education in Brazil.

The meta-frontier analysis suggests that municipalities that have between 5 and 50 thousand inhabitants are the ones that mostly make the efficiency frontier. This approach assumes that there is no difference in the education production function of the five groups of local authorities such that it would be possible for a municipality to emulate the best practices of another one that does not belong to the same cluster. According to this perspective, municipalities with up to five thousand inhabitants (22.3% of the sample) and the ones with more than five hundred thousand (0.8%) are the less efficient ones, using less than two-thirds of their resources efficiently.

In this sense, the comparison between the group and meta-frontiers shows that there is a technology gap between the local authorities. On the one hand, bigger municipalities might face more challenges to manage their school systems, which are likely to be more complex as they have more schools under their management. On the other hand, some small municipalities might have less qualified personnel in charge of their educational policies and are more likely to be captured by local elites.

In terms of policy implications, the efficiency analysis identifies the benchmark DMUs of each municipality that does not reach the frontier. That is, municipalities with efficiency scores lower than one can identify the local authorities that are somehow comparable to them and are on the frontier. Identifying the benchmark municipalities could contribute to the dialogue between municipal governments to share policy interventions aimed at increasing students' proficiency.

In addition to that, the efficiency scores of the municipalities could be used by the federal and state governments as an additional component when applying the rules for the distribution of educational resources. For example, the most important transfer from federal to municipal governments is made based on the number of inhabitants.[13] One possibility would be to award the municipalities based on their efficiency levels or/and how they are able to increase their efficiency over time. Plaček et al. (2020) conclude that DEA might be an effective accountability mechanism that could be used by policymakers and central governments to analyze policy results. Also, the authors claim that the efficiency scores could help citizens monitor their local governments, as the score interval between 0 and 1 seems to be quite easy to interpret.

This strategy is similar to the one employed by the state of Ceará, where the state government transfers the resources of the tax on goods and services (ICMS) based on education, health, and environmental indicators.[14] The advantage of the DEA approach would be to award municipalities based on how efficiently they are able to use their resources since there is the possibility that their educational outcomes have increased, but due to high levels of expenditure.

The fiscal space that would be achieved due to an increase in efficiency could be allocated to interventions to increase students' performance in a post-pandemic context where they are so much needed.

[13]Resources of the *Fundo de Participação dos Municípios (FPM).*
[14]Resources of the *Imposto Sobre Circulação de Mercadorias e Serviços (ICMS).*

References

Adda, J. (2016): "Economic activity and the spread of viral diseases: Evidence from high frequency data," *The Quarterly Journal of Economics*, 131, 891–941.

Afonso, A. and M. S. Aubyn (2005): "Non-parametric approaches to education and health efficiency in OECD countries," *Journal of applied economics*, 8, 227–246.

Ahamed, S. K., M. Naidu, and C. S. R. Reddy (2015): "Outliers in data envelopment analysis," *International Journal of Computer Science and Security (IJCSS)*, 9, 164–173.

Ahrens, A., C. B. Hansen, and M. Schaffer (2019): "LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation," .

Akresh, R., D. Halim, and M. Kleemans (2018): "Long-term and intergenerational effects of education: Evidence from school construction in Indonesia," Tech. rep., National Bureau of Economic Research.

Alexander, K. L., D. R. Entwisle, and L. S. Olson (2007): "Lasting consequences of the summer learning gap," *American sociological review*, 72, 167–180.

Allinder, R. M., L. S. Fuchs, D. Fuchs, and C. L. Hamlett (1992): "Effects of summer break on math and spelling performance as a function of grade level," *The Elementary School Journal*, 92, 451–460.

Alves, M. T. G., J. F. Soares, and F. P. Xavier (2016): "Desigualdades educacionais no ensino fundamental de 2005 a 2013: hiato entre grupos sociais," *Revista Brasileira de Sociologia*, 4, 49–82.

American School Counselor Association (2015): "Empirical research studies supporting the value of school counseling," .

Andersen, P. and N. C. Petersen (1993): "A procedure for ranking efficient units in data envelopment analysis," *Management science*, 39, 1261–1264.

Andrabi, T., B. Daniels, and J. Das (2020): "Human Capital Accumulation and Disasters: Evidence from the Pakistan Earthquake of 2005," *OSF. http://doi. org/10.17605/OSF. IO/3QG98*.

Angrist, J. D. and J.-S. Pischke (2008): *Mostly harmless econometrics*, Princeton university press.

Araújo Junior, J. N., W. R. Justo de, J. R. F. de Lima, M. d. O. FERREIRA, J. L. P. Araújo, and A. F. C. Pereira (2019): "Analysis on the Technical Efficiency of Northeast Municipal Expenditure with Basic Education: A DEA Approach and Malmquist's Index," *Embrapa Semiárido-Artigo em periódico indexado (ALICE).*

Athey, S. and G. Imbens (2006): "Identification and Inference in Nonlinear Difference-in-Differences Models," *Econometrica*, 74, 431–497.

Avkiran, N. K. (2011): "Association of DEA super-efficiency estimates with financial ratios: Investigating the case for Chinese banks," *Omega*, 39, 323–334.

Baker, M. (2013): "Industrial actions in schools: strikes and student achievement," *Canadian Journal of Economics/Revue canadienne d'économique*, 46, 1014–1036.

Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukerji, M. Shotland, and M. Walton (2017): "From proof of concept to scalable policies: Challenges and solutions, with an application," *Journal of Economic Perspectives*, 31, 73–102.

Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukherji, M. Shotland, and M. Walton (2016): "Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India," Tech. rep., National Bureau of Economic Research.

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007): "Remedying education: Evidence from two randomized experiments in India," *The Quarterly Journal of Economics*, 122, 1235–1264.

Banker, R. D. and H. Chang (2006): "The super-efficiency procedure for outlier identification, not for ranking efficient units," *European journal of operational research*, 175, 1311–1320.

Banker, R. D., H. Chang, and Z. Zheng (2017): "On the use of super-efficiency procedures for ranking efficient units and identifying outliers," *Annals of Operations Research*, 250, 21–35.

Banker, R. D. and J. L. Gifford (1988): "A relative efficiency model for the evaluation of public health nurse productivity," *Pittsburgh: Carnegie Mellon University.*

Belot, M. and D. Webbink (2010): "Do teacher strikes harm educational attainment of students?" *Labour*, 24, 391–406.

Boueri, R. (2015): "Modelos não paramétricos: Análise Envoltória de Dados (DEA)," in *Avaliação da Qualidade do Gasto Público e Mensuração da Eficiência*, ed. by R. Boueri, F. Rocha, and F. Rodopoulos, Brasília: Secretaria do Tesouro Nacional, chap. 8, 269–305.

Bryk, A. S. and S. W. Raudenbush (1989): "Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model," in *Multilevel analysis of educational data*, Elsevier, 159–204.

Charnes, A., W. W. Cooper, and E. Rhodes (1978): "Measuring the efficiency of decision making units," *European journal of operational research*, 2, 429–444.

Ciccone, A. and E. Papaioannou (2009): "Human capital, the structure of production, and growth," *The review of economics and statistics*, 91, 66–82.

Coleman, J. S. (1968): "Equality of educational opportunity," *Integrated education*, 6, 19–28.

Cooper, H., B. Nye, K. Charlton, J. Lindsay, and S. Greathouse (1996): "The effects of summer vacation on achievement test scores: A narrative and meta-analytic review," *Review of educational research*, 66, 227–268.

Crouch, L. and A. K. Gove (2011): "Leaps or one step at a time: Skirting or helping engage the debate? The case of reading," in *Policy debates in comparative, international, and development education*, Springer, 155–174.

Cunha, F. and J. Heckman (2007): "The technology of skill formation," *American Economic Review*, 97, 31–47.

Dawood, F. S., A. D. Iuliano, C. Reed, M. I. Meltzer, D. K. Shay, P.-Y. Cheng, D. Bandaranayake, R. F. Breiman, W. A. Brooks, P. Buchy, et al. (2012): "Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study," *The Lancet infectious diseases*, 12, 687–695.

Donnelly, R. and H. A. Patrinos (2021): "Learning loss during COVID-19: An early systematic review," *Prospects*, 1–9.

Duflo, A., J. Kiessel, and A. Lucas (2020): "Experimental Evidence on Alternative Policies to Increase Learning at Scale," Tech. rep., National Bureau of Economic Research.

Education at a Glance (2021): *Education at a Glance 2021*.

Faguet, J.-P. (1999): *Does decentralization increase responsiveness to local needs? Evidence from Bolivia*, The World Bank.

Filmer, D. (2002): *Autonomy, participation, and learning in Argentine schools: Findings and their implications for decentralization*, The World Bank.

Fryer Jr, R. G. and M. Howard-Noveck (2020): "High-dosage tutoring and reading achievement: evidence from New York City," *Journal of Labor Economics*, 38, 421–452.

Galiani, S., P. Gertler, and E. Schargrodsky (2008): "School decentralization: Helping the good get better, but leaving the poor behind," *Journal of public economics*, 92, 2106–2120.

Global Education Monitoring Report (2019): "Education as healing: Addressing the trauma of displacement through social and emotional learning," .

Gorard, S., B. H. See, and N. Siddiqui (2017): *The trials of evidence-based education: The promises, opportunities and problems of trials in education*, Routledge.

Haimovich, F., E. Vazquez, and M. Adelman (2021): "Scalable early warning systems for school dropout prevention: Evidence from a 4.000-school randomized controlled trial," Tech. rep., Documento de Trabajo.

Hanushek, E. A. (1981): "Throwing money at schools," *Journal of policy analysis and management*, 1, 19–41.

——— (1986): "The economics of schooling: Production and efficiency in public schools," *Journal of economic literature*, 24, 1141–1177.

——— (1996): "A more complete picture of school resource policies," *Review of educational research*, 66, 397–409.

——— (2006): "Alternative school policies and the benefits of general cognitive skills," *Economics of Education Review*, 25, 447–462.

Hanushek, E. A. and D. D. Kimko (2000): "Schooling, labor-force quality, and the growth of nations," *American economic review*, 90, 1184–1208.

Hanushek, E. A. and L. Woessmann (2012): "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation," *Journal of economic growth*, 17, 267–321.

Harris, D. N. and T. R. Sass (2009): "What Makes for a Good Teacher and Who Can Tell? Working Paper 30." *National Center for Analysis of Longitudinal Data in Education Research.*

Hernandez, D. J. (2011): "Double jeopardy: How third-grade reading skills and poverty influence high school graduation." *Annie E. Casey Foundation.*

Herrera, S. and G. Pang (2005): *Efficiency of public spending in developing countries: an efficiency frontier approach*, vol. 3645, World Bank Publications.

Jimenez, E. and Y. Sawada (1999): "Do community-managed schools work? An evaluation of El Salvador's EDUCO program," *The world bank economic review*, 13, 415–441.

King, E. and B. Ozler (2000): "What's decentralization got to do with learning? Endogenous school quality and student performance in Nicaragua," *Development Research Group, World Bank, Washington, DC.*

Klaiman, T., J. D. Kraemer, and M. A. Stoto (2011): "Variability in school closure decisions in response to 2009 H1N1: a qualitative systems improvement analysis," *BMC Public Health*, 11, 73.

Knoster, K. C. (2016): "Strategies for Addressing Student and Teacher Absenteeism: A Literature Review." *North Central Comprehensive Center.*

Kuhfeld, M., J. Soland, B. Tarasawa, A. Johnson, E. Ruzek, and J. Liu (2020): "Projecting the potential impact of COVID-19 school closures on academic achievement," *Educational Researcher*, 49, 549–565.

Lacey, M. and D. McNeil Jr (2009): "Fighting Deadly Flu, Mexico Shuts Schools," *The New York Times.*

Maldonado, J. and K. De Witte (2020): “The effect of school closures on standardised student test,” *FEB Research Report Department of Economics*.

Marcotte, D. E. and S. W. Hemelt (2008): “Unscheduled school closings and student performance,” *Education Finance and Policy*, 3, 316–338.

Mattos, E. and R. Terra (2015a): “Conceitos sobre eficiência,” in *Avaliação da Qualidade do Gasto Público e Mensuração da Eficiência*, ed. by R. Boueri, F. Rocha, and F. Rodopoulos, Brasília: Secretaria do Tesouro Nacional, chap. 6, 211–233.

——— (2015b): “Fundamentos microeconômicos da mensuração da ineficiência,” in *Avaliação da Qualidade do Gasto Público e Mensuração da Eficiência*, ed. by R. Boueri, F. Rocha, and F. Rodopoulos, Brasília: Secretaria do Tesouro Nacional, chap. 7, 235–266.

McCombs, J. S. (2011): *Making summer count: How summer programs can boost children's learning*, Rand Corporation.

McEwan, P. J. (2015): “Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments,” *Review of Educational Research*, 85, 353–394.

Melo Castro, C. R., G. da Silva de, M. E. Tannuri-Pianto, et al. (2017): “Education spending: more resources without management?” *Economia Aplicada*, 21, 285–309.

Menezes-Filho, N. A., L. Amaral, et al. (2009): “A relação entre gastos educacionais e desempenho escolar,” *São Paulo: Ibmec*.

Meyers, K. and M. A. Thomasson (2017): “Paralyzed by panic: Measuring the effect of school closures during the 1916 polio pandemic on educational attainment,” Tech. rep., National Bureau of Economic Research.

Muralidharan, K. and N. Prakash (2017): “Cycling to school: Increasing secondary school enrollment for girls in India,” *American Economic Journal: Applied Economics*, 9, 321–50.

Murnane, R. J. (1975): “The impact of school resources on the learning of inner city children.”
.

Oliveira, F. R., D. L. Casagrande, and G. Trevisan (2019): “Correção de fluxo e desempenho escolar: Uma avaliação de impacto dos programas Se Liga e Acelera,” Tech. rep., Faculdade de Economia da Universidade Federal do Mato Grosso.

Pigott, T. D., R. T. Williams, J. R. Polanin, and M.-J. Wu-Bohanon (2012): "Predicting Student Achievement with the Education Production-Function and Per-Pupil Expenditure: Synthesizing Regression Models from 1968-1994." *Society for Research on Educational Effectiveness.*

Plaček, M., M. Křápek, J. Čadil, and B. Hamerníková (2020): "The Influence of Excellence on Municipal Performance: Quasi-Experimental Evidence From the Czech Republic," *SAGE Open*, 10, 2158244020978232.

Rocha, F., J. Duarte, S. R. d. B. Gadelha, P. P. d. Oliveira, and L. F. V. N. Pereira (2013): "Texto para discussão n. 15: é possível atingir as metas para a educação sem aumentar os gastos? Uma análise para os municípios brasileiros," .

Rogers, F. H. and S. Sabarwal (2020): "The COVID-19 Pandemic: Shocks to Education and Policy Responses," Tech. rep., The World Bank.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb (2019): "Fast and wild: Bootstrap inference in Stata using boottest," *The Stata Journal*, 19, 4–60.

Ruggiero, J. (1998): "Non-discretionary inputs in data envelopment analysis," *European Journal of Operational Research*, 111, 461–469.

——— (2007): "A comparison of DEA and the stochastic frontier model using panel data," *International Transactions in Operational Research*, 14, 259–266.

Simar, L. and P. W. Wilson (1998): "Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models," *Management science*, 44, 49–61.

Somers, M.-A., W. Corrin, S. Sepanik, T. Salinger, J. Levin, and C. Zmach (2010): "The Enhanced Reading Opportunities Study Final Report: The Impact of Supplemental Literacy Courses for Struggling Ninth-Grade Readers. NCEE 2010-4021." *National Center for Education Evaluation and Regional Assistance.*

Sousa, M. d. C. S. d., F. Cribari-Neto, and B. D. Stosic (2005): "Explaining DEA technical efficiency scores in an outlier corrected environment: the case of public services in Brazilian municipalities," *Brazilian Review of Econometrics*, 25, 287–313.

Sutherland, D., R. Price, I. Joumard, and C. Nicq (2007): "Performance indicators for public spending efficiency in primary and secondary education," .

The World Health Organization (2010): "Pandemic (H1N1) 2009 - update 100," .

Thum, Y. M. and C. H. Hauser (2015): "NWEA 2015 MAP norms for student and school achievement status and growth," *Portland, OR: NWEA*.

Tomasik, M. J., L. A. Helbling, and U. Moser (2021): "Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland," *International Journal of Psychology*, 56, 566–576.

Tone, K. (2002): "A slacks-based measure of super-efficiency in data envelopment analysis," *European Journal of Operational Research*, 143, 32–41.

United Nations Development Group (2015): "Socio-economic impact of Ebola virus disease in West African countries," .

United Nations Development Programme (2015): "Assessing Sexual and Gender Based Violence during the Ebola Crisis in Sierra Leone," Tech. rep., UNPD.

Wongchai, A., W.-B. Liu, and K.-C. Peng (2012): "DEA metafrontier analysis on technical efficiency differences of national universities in Thailand," *International Journal on New Trends in Education and their implications*, 3, 3.

A Appendix to Chapter 1

Table A.1: Balance test: municipalities that extended the winter break versus municipalities that did not (2009)

	(1) Comparison Group		(2) Treatment Group		T-test Difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Average enrollment per school	629	316.33	12	531.22	-214.89***
		(7.52)		(62.36)	
Total enrollment	632	2,700.47	12	19,633.50	-16,933.03***
		(199.00)		(3,515.11)	
Teachers with tenure	565	0.39	12	0.60	-0.21**
		(0.01)		(0.10)	
H1N1 cases per 100,000 inhabitants	632	1.95	12	9.80	-7.85***
		(0.18)		(0.94)	
Population, in thousand	632	39.53	12	446.73	-407.19***
		(3.48)		(90.31)	
GDP per capita	610	26,116.72	12	50,073.24	-23,956.52***
		(524.71)		(6,472.35)	

Notes: The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. The municipality of São Paulo is not included in the Table since it had more than 11 million inhabitants in 2009. Therefore, its inclusion would distort the comparison between treatment and comparison municipalities. Source: *Prova Brasil*, Census of Education and IBGE.

Table A.2: Linear Regression of 2007 school characteristics on the decision to extend the winter break in 2009

	Extended the winter break
Math, 5th grade	-0.002**
	0.001
Portuguese, 5th grade	0.002*
	0.001
Retention - %	0.001
	0.001
Dropout - %	-0.006
	0.006
Class hours per day	-0.038***
	0.006
Students per teacher	0.000
	0.000
Lack of textbooks according to principals	0.000***
	0.000
Teachers classify the textbooks as great - %	0.000
	0.000
Teachers covered more than 80 percent of the curricula - %	0.000
	0.000
Principal managerial skills from teacher perspective - %	-0.093**
	0.032
Teacher with tenure - %	0.001***
	0.000
Student absenteeism as a big issue	0.001*
	0.000
Teacher absenteeism as a big issue	0.001**
	0.000
Deficit in learning is due to: students' low effort - %	-0.001***
	0.000
Deficit in learning is due to: students' bad behavior - %	0.000
	0.000
Students allocated into classrooms according to similar ages	0.000
	0.000
GDP per capita	0.000***
	0.000
Population	0.000***
	0.000
H1N1 cases per 100.000 inhabitants (Late July 2009)	0.000
	0.000
Constant	0.300***
	0.074
Obs	2963
Adj. R-squared	0.596

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. The Table shows the results of a linear regression at the school level, in which the dependent variable assumes the value of 1 if the school is located in a municipality that opted to extend children's winter break and 0, otherwise. The sample includes all municipalities of São Paulo in which the data on population, number of H1N1 confirmed cases, GDP per capita and education data are available (516 municipalities.

Table A.3: Fifth-graders' characteristics, locally-managed schools (2009)

	(1) Comparison Group		(2) Treatment Group		T-test Difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Proficiency in Math [SAEB scale 0 to 350]	2568	219.85 (0.40)	795	202.19 (0.56)	17.66***
Proficiency in Portuguese [SAEB scale 0 to 325]	2568	194.60 (0.33)	795	182.39 (0.52)	12.21***
Retention - %	2567	5.95 (0.12)	794	10.75 (0.30)	-4.80***
Dropout - %	2567	0.18 (0.01)	794	0.39 (0.03)	-0.22***
a Grade-promotion - %	2567	93.87 (0.12)	794	88.85 (0.30)	5.02***
Parents encourage to study - %	2564	97.93 (0.05)	795	96.93 (0.08)	1.01***
Parents encourage to do the homework - %	2564	96.74 (0.06)	795	95.50 (0.10)	1.24***
Parents encourage to read - %	2564	95.01 (0.07)	795	94.61 (0.10)	0.40***
Parents encourage to go to school - %	2564	97.14 (0.06)	795	95.70 (0.10)	1.44***
Parents talk about what happens in the school - %	2564	84.07 (0.14)	795	84.89 (0.18)	-0.81***
White students - %	2565	45.38 (0.27)	795	38.34 (0.29)	7.04***
Student lives with mother (or legal responsible) - %	2564	94.83 (0.07)	795	94.51 (0.10)	0.32**
Computer in the household - %	2565	49.80 (0.33)	795	58.43 (0.45)	-8.63***
Students' mother finished high school - %	2564	33.97 (0.30)	795	37.20 (0.46)	-3.23***
Student did preschool - %	2564	83.40 (0.20)	795	75.94 (0.29)	7.46***
Student has ever repeated and 0, otherwise - %	2564	21.55 (0.21)	795	20.53 (0.29)	1.02**
Student has ever dropped and 0, otherwise - %	2564	4.86 (0.08)	795	7.11 (0.15)	-2.25***
Student works - %	2564	11.86 (0.14)	795	13.27 (0.21)	-1.41***
Students per class	2568	27.44 (0.10)	795	31.41 (0.14)	-3.97***
Class hours per day	2568	5.02 (0.01)	795	4.79 (0.01)	0.22***
Insufficient performance in Math - %	2566	35.52 (0.32)	795	49.47 (0.49)	-13.95***
Insufficient performance in Portuguese - %	2566	56.87 (0.31)	795	66.07 (0.45)	-9.21***
Teacher always corrects Portuguese homework - %	2564	84.00 (0.22)	795	78.27 (0.33)	5.74***
Teacher always corrects Math homework - %	2564	85.58 (0.21)	795	80.91 (0.30)	4.67***
Schools with computer lab - %	2568	58.68 (0.97)	795	93.58 (0.87)	-34.90***
Schools with science lab - %	2568	5.45 (0.45)	795	20.00 (1.42)	-14.55***
Schools sport court - %	2568	66.51 (0.93)	795	90.44 (1.04)	-23.93***
Schools with library - %	2568	24.65 (0.85)	795	23.40 (1.50)	1.25
Schools internet access - %	2568	90.50 (0.58)	795	93.71 (0.86)	-3.21***
Length of the school year (days)	2568	318.76 (0.21)	795	313.95 (0.53)	4.80***
Total enrollment, all grades	2568	486.22 (5.46)	795	872.01 (12.41)	-385.79***
GDP per capita of the municipality, in 2019 BRL	2440	30,945.36 (314.88)	795	57,707.66 (430.57)	-26,762.30***

Notes: Locally-managed schools in São Paulo. The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.4: Fifth-graders' characteristics, state and locally-managed schools in $G = 1$ (2009)

Variable	(1) Locally-managed N	Mean/SE	(2) State-managed N	Mean/SE	T-test Difference (1)-(2)
Proficiency in Math [SAEB scale 0 to 350]	1334	216.95 (0.54)	929	215.01 (0.62)	1.94**
Proficiency in Portuguese [SAEB scale 0 to 325]	1334	193.39 (0.46)	929	189.67 (0.52)	3.72***
Retention - %	1334	5.69 (0.16)	929	6.10 (0.19)	-0.41*
Dropout - %	1334	0.21 (0.02)	929	0.17 (0.02)	0.04
Grade-promotion - %	1334	94.10 (0.16)	929	93.73 (0.19)	0.37
Parents encourage to study - %	1333	97.92 (0.06)	929	97.80 (0.08)	0.12
Parents encourage to do the homework - %	1333	96.72 (0.08)	929	96.36 (0.11)	0.36***
Parents encourage to read - %	1333	94.96 (0.09)	929	95.62 (0.10)	-0.66***
Parents encourage to go to school - %	1333	97.15 (0.07)	929	96.81 (0.09)	0.34***
Parents talk about what happens in the school - %	1333	83.36 (0.19)	929	85.08 (0.20)	-1.72***
White students - %	1333	44.18 (0.36)	929	44.76 (0.43)	-0.58
Student lives with mother (or legal responsible) - %	1333	95.11 (0.09)	929	94.95 (0.10)	0.16
Computer in the household - %	1333	53.27 (0.45)	929	52.39 (0.53)	0.88
Students' mother finished high school - %	1333	36.09 (0.43)	929	35.75 (0.49)	0.34
Student did preschool - %	1333	82.26 (0.28)	929	81.52 (0.33)	0.74*
Student has ever repeated and 0, otherwise - %	1333	19.56 (0.28)	929	16.40 (0.30)	3.16***
Student has ever dropped and 0, otherwise - %	1333	4.90 (0.11)	929	5.79 (0.14)	-0.89***
Student works - %	1333	11.15 (0.17)	929	12.66 (0.23)	-1.51***
Students per class	1334	28.71 (0.14)	929	28.82 (0.16)	-0.10
Class hours per day	1334	4.91 (0.02)	929	5.31 (0.03)	-0.40***
Insufficient performance in Math - %	1334	37.59 (0.44)	929	40.02 (0.50)	-2.43***
Insufficient performance in Portuguese - %	1334	57.73 (0.43)	929	60.46 (0.48)	-2.73***
Teacher always corrects Portuguese homework - %	1333	81.95 (0.31)	929	82.10 (0.36)	-0.16
Teacher always corrects Math homework - %	1333	83.62 (0.30)	929	83.89 (0.34)	-0.27
Schools with computer lab - %	1334	57.80 (1.35)	929	92.25 (0.88)	-34.45***
Schools with science lab - %	1334	6.90 (0.69)	929	6.89 (0.83)	0.01
Schools sport court - %	1334	61.69 (1.33)	929	84.39 (1.19)	-22.70***
Schools with library - %	1334	25.86 (1.20)	929	2.91 (0.55)	22.96***
Schools internet access - %	1334	88.46 (0.88)	929	98.28 (0.43)	-9.82***
Length of the school year (days)	1334	320.22 (0.34)	929	315.01 (0.25)	5.21***
Total enrollment, all grades	1334	541.53 (8.28)	929	638.05 (12.99)	-96.52***
GDP per capita of the municipality, in 2019 BRL	1288	33,284.83 (370.67)	914	33,830.75 (434.43)	-545.92

Notes: State and locally-managed schools in São Paulo. Group of municipalities in $G = 1$ (Table 1.2). The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.5: Fifth-graders' characteristics, state and locally-managed schools in $G = 0$ (2009)

	(1)		(2)		T-test
	Locally-managed		State-managed		Difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Proficiency in Math [SAEB scale 0 to 350]	759	201.11	868	210.28	-9.16***
		(0.58)		(0.54)	
Proficiency in Portuguese [SAEB scale 0 to 325]	759	181.48	868	188.01	-6.53***
		(0.53)		(0.48)	
Retention - %	758	11.40	868	5.78	5.62***
		(0.30)		(0.18)	
Dropout - %	758	0.40	868	0.37	0.03
		(0.03)		(0.03)	
Grade-promotion - %	758	88.20	868	93.85	-5.65***
		(0.31)		(0.19)	
Parents encourage to study - %	759	96.81	867	97.58	-0.77***
		(0.09)		(0.08)	
Parents encourage to do the homework - %	759	95.43	867	96.22	-0.79***
		(0.10)		(0.12)	
Parents encourage to read - %	759	94.64	867	95.38	-0.74***
		(0.10)		(0.10)	
Parents encourage to go to school - %	759	95.64	867	96.40	-0.76***
		(0.11)		(0.09)	
Parents talk about what happens in the school - %	759	85.10	867	85.08	0.02
		(0.18)		(0.16)	
White students - %	759	38.02	867	40.36	-2.34***
		(0.29)		(0.32)	
Student lives with mother (or legal responsible) - %	759	94.42	867	94.78	-0.36***
		(0.11)		(0.09)	
Computer in the household - %	759	58.81	867	61.20	-2.39***
		(0.43)		(0.44)	
Students' mother finished high school - %	759	36.59	867	39.63	-3.04***
		(0.43)		(0.45)	
Student did preschool - %	759	75.21	867	79.90	-4.70***
		(0.29)		(0.30)	
Student has ever repeated and 0, otherwise - %	759	21.16	867	14.98	6.18***
		(0.32)		(0.26)	
Student has ever dropped and 0, otherwise - %	759	7.40	867	6.42	0.98***
		(0.16)		(0.14)	
Student works - %	759	13.42	867	12.09	1.33***
		(0.22)		(0.17)	
Students per class	759	31.76	868	31.37	0.39*
		(0.14)		(0.14)	
Class hours per day	759	4.74	868	5.17	-0.43***
		(0.02)		(0.02)	
Insufficient performance in Math - %	759	50.41	868	42.96	7.45***
		(0.51)		(0.46)	
Insufficient performance in Portuguese - %	759	66.70	868	61.64	5.05***
		(0.47)		(0.43)	
Teacher always corrects Portuguese homework - %	759	78.19	867	80.30	-2.12***
		(0.35)		(0.31)	
Teacher always corrects Math homework - %	759	80.86	867	82.43	-1.56***
		(0.32)		(0.30)	
Schools with computer lab - %	759	94.07	868	89.06	5.02***
		(0.86)		(1.06)	
Schools with science lab - %	759	20.95	868	11.64	9.31***
		(1.48)		(1.09)	
Schools sport court - %	759	89.99	868	85.83	4.16**
		(1.09)		(1.18)	
Schools with library - %	759	20.95	868	3.80	17.15***
		(1.48)		(0.65)	
Schools internet access - %	759	93.68	868	97.47	-3.79***
		(0.88)		(0.53)	
Length of the school year (days)	759	313.74	868	313.50	0.24
		(0.56)		(0.28)	
Total enrollment, all grades	759	888.28	868	854.12	34.16*
		(12.22)		(15.19)	
GDP per capita of the municipality, in 2019 BRL	759	58,012.21	868	56,747.25	1,264.96***
		(384.71)		(306.31)	

Notes: State and locally-managed schools in São Paulo. Group of municipalities in $G = 0$ (Table 1.2). The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.6: Fifth-graders' characteristics, state-managed schools (2009)

Variable	(1) State-managed in G = 0 N	Mean/SE	(2) State-managed in G = 1 N	Mean/SE	T-test Difference (1)-(2)
Proficiency in Math [SAEB scale 0 to 350]	868	210.28 (0.54)	929	215.01 (0.62)	-4.73***
Proficiency in Portuguese [SAEB scale 0 to 325]	868	188.01 (0.48)	929	189.67 (0.52)	-1.66**
Retention - %	868	5.78 (0.18)	929	6.10 (0.19)	-0.32
Dropout - %	868	0.37 (0.03)	929	0.17 (0.02)	0.20***
Grade-promotion - %	868	93.85 (0.19)	929	93.73 (0.19)	0.12
Parents encourage to study - %	867	97.58 (0.08)	929	97.80 (0.08)	-0.22**
Parents encourage to do the homework - %	867	96.22 (0.12)	929	96.36 (0.11)	-0.14
Parents encourage to read - %	867	95.38 (0.10)	929	95.62 (0.10)	-0.24*
Parents encourage to go to school - %	867	96.40 (0.09)	929	96.81 (0.09)	-0.41***
Parents talk about what happens in the school - %	867	85.08 (0.16)	929	85.08 (0.20)	-0.00
White students - %	867	40.36 (0.32)	929	44.76 (0.43)	-4.40***
Student lives with mother (or legal responsible) - %	867	94.78 (0.09)	929	94.95 (0.10)	-0.17
Computer in the household - %	867	61.20 (0.44)	929	52.39 (0.53)	8.81***
Students' mother finished high school - %	867	39.63 (0.45)	929	35.75 (0.49)	3.88***
Student did preschool - %	867	79.90 (0.30)	929	81.52 (0.33)	-1.62***
Student has ever repeated and 0, otherwise - %	867	14.98 (0.26)	929	16.40 (0.30)	-1.42***
Student has ever dropped and 0, otherwise - %	867	6.42 (0.14)	929	5.79 (0.14)	0.64***
Student works - %	867	12.09 (0.17)	929	12.66 (0.23)	-0.57**
Students per class	868	31.37 (0.14)	929	28.82 (0.16)	2.55***
Class hours per day	868	5.17 (0.02)	929	5.31 (0.03)	-0.14***
Insufficient performance in Math - %	868	42.96 (0.46)	929	40.02 (0.50)	2.93***
Insufficient performance in Portuguese - %	868	61.64 (0.43)	929	60.46 (0.48)	1.18*
Teacher always corrects Portuguese homework - %	867	80.30 (0.31)	929	82.10 (0.36)	-1.80***
Teacher always corrects Math homework - %	867	82.43 (0.30)	929	83.89 (0.34)	-1.46***
Schools with computer lab - %	868	89.06 (1.06)	929	92.25 (0.88)	-3.19**
Schools with science lab - %	868	11.64 (1.09)	929	6.89 (0.83)	4.75***
Schools sport court - %	868	85.83 (1.18)	929	84.39 (1.19)	1.44
Schools with library - %	868	3.80 (0.65)	929	2.91 (0.55)	0.90
Schools internet access - %	868	97.47 (0.53)	929	98.28 (0.43)	-0.81
Length of the school year (days)	868	313.50 (0.28)	929	315.01 (0.25)	-1.52***
Total enrollment, all grades	868	854.12 (15.19)	929	638.05 (12.99)	216.07***
GDP per capita of the municipality, in 2019 BRL	868	56,747.25 (306.31)	914	33,830.75 (434.43)	22,916.50***

Notes: State-managed schools in São Paulo. The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.7: Teachers' and principals' characteristics, locally-managed schools (2009)

Variable	(1) Comparison Group N	Mean/SE	(2) Treatment Group N	Mean/SE	T-test Difference (1)-(2)
Teacher with tenure - %	2348	48.73 (0.87)	719	67.44 (1.21)	-18.71***
Teacher with less than 40 years old - %	2353	48.53 (0.75)	722	38.43 (1.17)	10.10***
Principal managerial skills from teacher perspective	2253	0.76 (0.00)	643	0.72 (0.01)	0.04***
Index for the violence the teacher faces in the school	2344	0.14 (0.01)	716	0.28 (0.01)	-0.14***
Teacher expects that almost all students will finish 9th grade - %	2326	88.76 (0.48)	721	87.36 (0.78)	1.41
Teacher expects that almost all students will finish high school - %	2312	60.72 (0.76)	715	53.35 (1.26)	7.37***
Teachers covered more than 80 percent of the curricula - %	2351	54.67 (0.78)	722	45.16 (1.23)	9.51***
Teachers always participate of the work decisions - %	2236	56.40 (0.80)	659	54.45 (1.38)	1.95
Teachers say that all the students have textbooks - %	2154	70.75 (0.83)	644	64.69 (1.45)	6.06***
Teachers classify the textbooks as great - %	2178	18.69 (0.68)	667	15.17 (1.07)	3.52**
Teachers' salary is less than 3 minimum wage - %	1980	33.07 (0.89)	665	13.49 (0.99)	19.58***
Deficit in learning is due to: students' low effort - %	2348	81.72 (0.59)	721	76.33 (0.98)	5.39***
Deficit in learning is due to: students' absenteeism - %	2210	31.76 (0.91)	650	36.09 (1.61)	-4.33**
Deficit in learning is due to: students' bad behavior - %	2345	56.91 (0.76)	718	61.24 (1.16)	-4.34***
Teachers with the correct degree to teach Portuguese - %	2133	69.39 (0.78)	662	74.10 (1.52)	-4.71***
Teachers with the correct degree to teach Math - %	2133	63.32 (0.81)	662	69.54 (1.59)	-6.22***
Principal has organized Teachers' training last two years, %	1716	56.99 (1.20)	557	64.63 (2.03)	-7.64***
Lack of textbooks according to principals, %	1671	29.20 (1.11)	553	51.72 (2.13)	-22.51***
Principal was appointed for the position, %	1731	47.66 (1.20)	565	14.51 (1.48)	33.15***
Teacher absenteeism as a big issue, %	1744	8.08 (0.65)	561	26.38 (1.86)	-18.30***
Student absenteeism as a big issue, %	1746	4.41 (0.49)	566	8.30 (1.16)	-3.89***
Students allocated into classrooms according to similar age, %	1660	35.00 (1.17)	549	27.69 (1.91)	7.31***
Students allocated into classrooms according to hetero. performance, %	1660	41.08 (1.21)	549	60.29 (2.09)	-19.21***

Notes: Locally-managed schools in São Paulo. The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.8: Teachers' and principals' characteristics, state and locally-managed schools in $G = 1$ (2009)

Variable	(1) Locally-managed N	(1) Locally-managed Mean/SE	(2) State-managed N	(2) State-managed Mean/SE	T-test Difference (1)-(2)
Teacher with tenure - %	1220	45.65	852	57.22	-11.57***
		(1.24)		(1.22)	
Teacher with less than 40 years old - %	1224	52.24	852	24.83	27.41***
		(1.02)		(1.04)	
Principal managerial skills from teacher perspective	1170	0.74	809	0.78	-0.04***
		(0.00)		(0.01)	
Index for the violence the teacher faces in the school	1220	0.15	851	0.12	0.03**
		(0.01)		(0.01)	
Teacher expects that almost all students will finish 9th grade - %	1212	87.37	850	89.54	-2.16**
		(0.69)		(0.71)	
Teacher expects that almost all students will finish high school - %	1203	56.83	843	66.06	-9.23***
		(1.06)		(1.19)	
Teachers covered more than 80 percent of the curricula - %	1222	51.13	852	51.57	-0.45
		(1.05)		(1.24)	
Teachers always participate of the work decisions - %	1170	54.84	810	59.90	-5.06***
		(1.09)		(1.24)	
Teachers say that all the students have textbooks - %	1133	64.71	787	76.00	-11.29***
		(1.19)		(1.21)	
Teachers classify the textbooks as great - %	1126	15.10	791	24.19	-9.09***
		(0.84)		(1.15)	
Teachers' salary is less than 3 minimum wage - %	1073	25.75	731	26.34	-0.59
		(1.10)		(1.30)	
Deficit in learning is due to: students' low effort - %	1219	80.67	853	77.90	2.77**
		(0.82)		(1.00)	
Deficit in learning is due to: students' absenteeism - %	1153	37.31	808	35.41	1.90
		(1.32)		(1.59)	
Deficit in learning is due to: students' bad behavior - %	1216	58.82	851	54.25	4.57***
		(1.02)		(1.23)	
Teachers with the correct degree to teach Portuguese - %	1117	71.88	748	64.38	7.50***
		(1.01)		(1.26)	
Teachers with the correct degree to teach Math - %	1117	64.75	748	57.92	6.83***
		(1.07)		(1.31)	
Principal has organized Teachers' training last two years, %	911	60.26	865	65.55	-5.29**
		(1.62)		(1.62)	
Lack of textbooks according to principals, %	891	34.01	869	25.66	8.35***
		(1.59)		(1.48)	
Principal was appointed for the position, %	927	41.75	869	8.17	33.58***
		(1.62)		(0.93)	
Teacher absenteeism as a big issue, %	928	8.41	871	19.75	-11.34***
		(0.91)		(1.35)	
Student absenteeism as a big issue, %	927	5.39	874	6.98	-1.59
		(0.74)		(0.86)	
Students allocated into classrooms according to similar age, %	876	39.16	855	36.96	2.20
		(1.65)		(1.65)	
Students allocated into classrooms according to hetero. performance, %	876	39.50	855	38.60	0.90
		(1.65)		(1.67)	

Notes: *Prova Brasil* and Census of Education. State and locally-managed schools in São Paulo. Group of municipalities in $G = 1$ (Table 1.2). The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Table A.9: Teachers' and principals' characteristics, state and locally-managed schools in $G = 0$ (2009)

Variable	(1) Locally-managed N	(1) Locally-managed Mean/SE	(2) State-managed N	(2) State-managed Mean/SE	T-test Difference (1)-(2)
Teacher with tenure - %	681	69.63 (1.25)	772	50.81 (1.18)	18.82***
Teacher with less than 40 years old - %	686	36.58 (1.18)	773	27.70 (0.94)	8.88***
Principal managerial skills from teacher perspective	608	0.72 (0.01)	709	0.75 (0.01)	-0.03***
Index for the violence the teacher faces in the school	680	0.29 (0.01)	769	0.26 (0.01)	0.03
Teacher expects that almost all students will finish 9th grade - %	685	86.80 (0.83)	772	88.20 (0.69)	-1.40
Teacher expects that almost all students will finish high school - %	680	51.58 (1.29)	768	64.85 (1.07)	-13.27***
Teachers covered more than 80 percent of the curricula - %	686	42.18 (1.26)	773	51.28 (1.13)	-9.10***
Teachers always participate of the work decisions - %	625	53.71 (1.43)	722	57.69 (1.24)	-3.98**
Teachers say that all the students have textbooks - %	606	64.71 (1.49)	683	67.35 (1.31)	-2.64
Teachers classify the textbooks as great - %	630	15.19 (1.10)	729	21.32 (1.11)	-6.13***
Teachers' salary is less than 3 minimum wage - %	625	12.50 (0.97)	723	24.54 (1.14)	-12.03***
Deficit in learning is due to: students' low effort - %	685	76.95 (1.01)	773	73.02 (1.00)	3.93***
Deficit in learning is due to: students' absenteeism - %	614	38.16 (1.70)	707	35.00 (1.55)	3.17
Deficit in learning is due to: students' bad behavior - %	682	63.12 (1.19)	770	55.03 (1.09)	8.09***
Teachers with the correct degree to teach Portuguese - %	626	73.07 (1.59)	667	66.15 (1.22)	6.93***
Teachers with the correct degree to teach Math - %	626	68.44 (1.66)	667	61.63 (1.26)	6.81***
Principal has organized Teachers' training last two years, %	540	65.19 (2.05)	794	62.97 (1.71)	2.21
Lack of textbooks according to principals, %	534	53.00 (2.16)	801	26.72 (1.56)	26.28***
Principal was appointed for the position, %	548	15.69 (1.56)	797	7.28 (0.92)	8.42***
Teacher absenteeism as a big issue, %	542	29.52 (1.96)	802	18.45 (1.37)	11.07***
Student absenteeism as a big issue, %	549	9.11 (1.23)	805	6.96 (0.90)	2.15
Students allocated into classrooms according to similar age, %	530	27.92 (1.95)	773	36.61 (1.73)	-8.69***
Students allocated into classrooms according to hetero. performance, %	530	60.57 (2.12)	773	42.69 (1.78)	17.88***

Notes: State and locally-managed schools in São Paulo. Group of municipalities in $G = 0$ (Table 1.2). The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.10: Teachers' and principals' characteristics, state-managed schools (2009)

Variable	(1) State-managed, G=0 N	(1) State-managed, G=0 Mean/SE	(2) State-managed, G=1 N	(2) State-managed, G=1 Mean/SE	T-test Difference (1)-(2)
Teacher with tenure - %	772	50.81 (1.18)	852	57.22 (1.22)	-6.41***
Teacher with less than 40 years old - %	773	27.70 (0.94)	852	24.83 (1.04)	2.87**
Principal managerial skills from teacher perspective	709	0.75 (0.01)	809	0.78 (0.01)	-0.04***
Index for the violence the teacher faces in the school	769	0.26 (0.01)	851	0.12 (0.01)	0.14***
Teacher expects that almost all students will finish 9th grade - %	772	88.20 (0.69)	850	89.54 (0.71)	-1.34
Teacher expects that almost all students will finish high school - %	768	64.85 (1.07)	843	66.06 (1.19)	-1.22
Teachers covered more than 80 percent of the curricula - %	773	51.28 (1.13)	852	51.57 (1.24)	-0.29
Teachers always participate of the work decisions - %	722	57.69 (1.24)	810	59.90 (1.24)	-2.21
Teachers say that all the students have textbooks - %	683	67.35 (1.31)	787	76.00 (1.21)	-8.65***
Teachers classify the textbooks as great - %	729	21.32 (1.11)	791	24.19 (1.15)	-2.87*
Teachers' salary is less than 3 minimum wage - %	723	24.54 (1.14)	731	26.34 (1.30)	-1.80
Deficit in learning is due to: students' low effort - %	773	73.02 (1.00)	853	77.90 (1.00)	-4.88***
Deficit in learning is due to: students' absenteeism - %	707	35.00 (1.55)	808	35.41 (1.59)	-0.41
Deficit in learning is due to: students' bad behavior - %	770	55.03 (1.09)	851	54.25 (1.23)	0.78
Teachers with the correct degree to teach Portuguese - %	667	66.15 (1.22)	748	64.38 (1.26)	1.77
Teachers with the correct degree to teach Math - %	667	61.63 (1.26)	748	57.92 (1.31)	3.71**
Principal has organized Teachers' training last two years, %	794	62.97 (1.71)	865	65.55 (1.62)	-2.58
Lack of textbooks according to principals, %	801	26.72 (1.56)	869	25.66 (1.48)	1.05
Principal was appointed for the position, %	797	7.28 (0.92)	869	8.17 (0.93)	-0.89
Teacher absenteeism as a big issue, %	802	18.45 (1.37)	871	19.75 (1.35)	-1.29
Student absenteeism as a big issue, %	805	6.96 (0.90)	874	6.98 (0.86)	-0.02
Students allocated into classrooms according to similar age, %	773	36.61 (1.73)	855	36.96 (1.65)	-0.35
Students allocated into classrooms according to hetero. performance, %	773	42.69 (1.78)	855	38.60 (1.67)	4.09*

Notes: State-managed schools in São Paulo. The value displayed for t-tests is the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. Source: *Prova Brasil* and Census of Education.

Table A.11: Placebo test for students' performance in treatment and comparison groups (2005-2007)

	Fifth-graders				Ninth-graders			
	Math		Portuguese		Math		Portuguese	
	I	II	I	II	I	II	I	II
State-managed	-4.564**	3.873	-2.315	5.804	-3.199	-4.708	-0.578	1.783
	1.72	6.31	1.49	5.5	2.16	9.61	1.96	8.69
Post-treatment year (2007)	12.482***	14.960*	-0.166	2.426	3.422	5.269	6.057**	11.212
	1.68	6.18	1.45	5.38	2.18	9.99	1.98	9.03
State-managed × post-treatment	-6.250**	-10.219	-5.844**	-11.406	-2.917	-6.937	-2.914	-10.515
	2.38	8.71	2.06	7.58	2.82	13.18	2.56	11.91
G = 1		9.08		6.78		-3.204		-2.634
		4.74		4.13		7.65		6.91
State-managed × G = 1		-8.437		-8.119		1.509		-2.362
		6.54		5.69		9.85		8.9
Post-treatment × with G = 1		-2.478		-2.592		-1.847		-5.155
		6.39		5.57		10.22		9.24
Triple DiD		3.969		5.562		4.02		7.601
		9.02		7.85		13.48		12.18
N	554	594	554	594	485	506	485	506
r2_a	0.171	0.174	0.067	0.066	0.027	0.023	0.031	0.03

Notes: The authors' estimate is based on data from *Prova Brasil*. The regressions are run at the municipality level for the years 2005 and 2007 as data at the school level is only available for the state-managed network since 2007. Sample of municipalities with at least one state and one locally-managed school. Columns I show the results of the following equation: $y_{mt} = \beta_0 + \beta_1 E + \beta_2 D + \beta_3 E \times D + v_{mt}$. Columns II show the results of the following equation: $y_{mt} = \beta_0 + \beta_1 E + \beta_2 D + \beta_3 E \times D + \beta_4 G + \beta_5 E \times G + \beta_6 G \times D + \beta_7 E \times G \times D + v_{mt}$. $E = 1$ for state-managed schools and 0 otherwise. $G = 1$ for the 112 municipalities whose local authorities did not extend children's winter break. $G = 0$ for the group of 10 municipalities whose local authorities extended children's winter break (Table 1.2). $D = 1$ for 2007 and $D = 0$ for 2005. β_3 is the DiD estimate and β_7 is the triple DiD estimate. For primary education, β_3 is statistically significant for both math and Portuguese. Therefore, even before the school closures, I do not have evidence that locally and state-managed schools have parallel trends. β_7 is not statistically significant, meaning that, before the shutdowns, the variation of students' proficiency in state and locally-managed schools in $G = 1$ minus the variation of students' proficiency in state and locally-managed schools in $G = 0$ is not statistically significant. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Table A.12: Placebo test for grade-promotion, retention, and dropout (2007-2008)

	Fifth-graders						Ninth-graders					
	Grade-promotion		Retention		Dropout		Grade-promotion		Retention		Dropout	
	I	II	I	II	I	II	I	II	I	II	I	II
State-managed	-2.489***	5.133***	2.439***	-5.162***	0.05	0.03	-9.896***	-5.514***	7.910***	4.315***	1.986***	1.199***
	0.27	0.37	0.27	0.37	0.04	0.07	0.55	0.57	0.51	0.53	0.22	0.22
Post-treatment year (2008)	0.618**	2.196***	-0.583*	-2.042***	-0.035	-0.154*	0.442	4.298***	-0.208	-4.151***	-0.234	-0.147
	0.24	0.39	0.24	0.38	0.04	0.07	0.62	0.62	0.57	0.58	0.25	0.24
State-managed × post-treatment	0.820*	0.455	-0.774*	-0.498	-0.046	0.043	1.864*	-1.185	-1.547*	1.418	-0.317	-0.232
	0.37	0.52	0.36	0.51	0.05	0.1	0.74	0.8	0.68	0.74	0.3	0.31
G = 1		1.901		-1.572		-0.329		14.575***		-11.341**		-3.234*
		2.09		2.04		0.38		3.88		3.59		1.51
State-managed × G = 1		-7.622***		7.601***		0.02		-4.382***		3.595***		0.787*
		0.48		0.47		0.09		0.83		0.77		0.32
Post-treatment × with G = 1		-1.577***		1.459**		0.119		-3.855***		3.942***		-0.087
		0.47		0.46		0.09		0.91		0.85		0.36
Triple DiD		0.365		-0.277		-0.089		3.049**		-2.965**		-0.084
		0.66		0.65		0.12		1.13		1.05		0.44
N	5275	8487	5275	8487	5275	8487	3161	5841	3161	5841	3161	5841
r2_a	0.167	0.196	0.167	0.195	0.04	0.016	0.271	0.244	0.234	0.217	0.145	0.089

Notes: The authors' estimate is based on data from *Prova Brasil*. G as defined in the empirical strategy (Table 1.2). The regressions are run at the school level for the years 2007 and 2008. Sample of municipalities with at least one state and one locally-managed school. Columns I show the results of the following equation: $y_{mt} = \beta_0 + \beta_1 E + \beta_2 D + \beta_3 E \times D + \upsilon_{mt}$. Columns II show the results of the following equation: $y_{mt} = \beta_0 + \beta_1 E + \beta_2 D + \beta_3 E \times D + \beta_4 G + \beta_5 E \times G + \beta_6 G \times D + \beta_7 E \times G \times D + \upsilon_{mt}$. $E = 1$ for state-managed schools and 0 otherwise. $G = 1$ for the 112 municipalities whose local authorities did not extend children's winter break. $G = 0$ for the group of 10 municipalities whose local authorities extended children's winter break (Table 1.2). $D = 1$ for 2007 and $D = 0$ for 2005. β_3 is the DiD estimate and β_7 is the triple DiD estimate. For primary education, β_3 is statistically significant in all DiD specifications. Therefore, even before the school closures, I do not have evidence that locally and state-managed schools have parallel trends. For primary education, β_7 is not statistically significant, meaning that, before the shutdowns, the variation of students' outcomes in state and locally-managed schools in $G = 1$ minus the variation of students' outcomes (grade promotion, retention, and dropout) in state and locally-managed schools in $G = 0$ is not statistically significant. Coefficients are converted to percentages (multiplied by 100). ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels.

Table A.13: Impact of the school shutdowns on students' learning (with placebo), fifth-grade (2005-2009)

	Math					Portuguese				
	DiD	DiD	DiD	Triple D	Triple D	DiD	DiD	DiD	Triple D	Triple D
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Placebo	0.27					0.8				
	(0.97)					(0.89)				
Wild-bootstrap p-value	0.80					0.53				
H1N1		-3.26**	-2.75**	-4.56***	-4.18***		-0.76	-0.54	-3.69***	-4.18***
		(1.18)	(1.23)	(0.85)	(1.28)		(0.86)	(0.95)	(0.73)	(1.28)
Wild-bootstrap p-value		0.04	0.04	0.00	0.00		0.39	0.61	0.00	0.00
N. schools	5291	3912	3912	5164	5164	5291	3912	3912	5164	5164
Adj. R-squared	0.6	0.8	0.8	0.8	0.7	0.5	0.8	0.8	0.8	0.7
Specifications										
(A) Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(B) Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(C) % teachers in both networks	No	No	Yes	No	Yes	No	No	Yes	No	Yes
(D) % teachers management	No	No	Yes	No	Yes	No	No	Yes	No	Yes
(E) School FE	No	No	No	No	Yes	No	No	No	No	Yes

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the managerial practices intervention. (E) schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented the managerial practices intervention. The row Placebo shows the estimates in which the pre-treatment year is 2005 and the post-treatment year is 2007.

Table A.14: Impact of school shutdowns on grade-promotion, retention and dropout, fifth-grade

	Grade-promotion				Retention				Dropout			
	Triple D	DiD	Triple D	Triple D	Triple D	DiD	Triple D	Triple D	Triple D	DiD	Triple D	Triple D
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Placebo	0.413				-0.180				-0.220**			
	(0.58)				(0.58)				(0.10)			
p-value	0.479				0.757				0.032			
H1N1		1.050*	0.934	-0.139		-1.263**	-1.179	-0.281		-0.151*	-0.161*	0.02
		(0.61)	(0.83)	(0.83)		(0.63)	(0.85)	(0.83)		(0.08)	(0.09)	(0.10)
p-value		0.085	0.264	0.867		0.046	0.167	0.736		0.074	0.075	0.839
N. schools	7,914	3,836	5,066	5,066	7,914	4,029	5,264	5,264	7,914	5,477	6,475	6,475
R2	0.30	0.50	0.40	0.10	0.30	0.50	0.40	0.10	0.00	0.00	0.00	0.00
Treatment Group before the school shutdowns (2007)												
Mean	90.62	86.37	90.93	90.93	9.07	13.49	8.78	8.78	0.31	0.49	0.29	0.29
Standard Deviation	9.5	11.28	6.88	6.88	9.36	10.94	6.76	6.76	1.71	3.29	0.81	0.81
Estimate of ATT (in sd)	0.04	0.09	0.14	-0.02	-0.02	-0.12	-0.17	-0.04	-0.13	-0.05	-0.20	0.03
Specifications												
(A) Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(B) Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(C) % teachers in both networks	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes
(D) % teachers management	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes
(E) School FE	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. ATT regressions coefficients are converted to percentages (multiplied by 100). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the managerial practices intervention. (E) schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented the managerial practices intervention. The row Placebo shows the estimates in which the pre-treatment year is 2007 and the post-treatment year is 2008.

Table A.15: Impact of the school shutdowns on students' learning (with and without schools with managerial practices intervention), fifth-grade (2007-2009)

Estimated decrease in Math and Portuguese Proficiency, SAEB scale										
Excluding state-managed schools that participated of the managerial practices intervention										
	Math					**Portuguese**				
	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D
	(4)	(5)	(6)	(7)	(8)	(4)	(5)	(6)	(7)	(8)
H1N1	-4.56***	-4.47***	-4.31***	-4.26***	-4.18***	-3.69***	-3.61***	-3.63***	-4.26***	-2.87***
	(0.85)	(0.90)	(0.91)	(0.95)	(1.28)	(0.73)	(0.76)	(0.75)	(0.95)	(1.05)
p-value	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.007
95% CI	[-6.3,-2.9]	[-6.3,-2.7]	[-6.1,-2.5]	[-6.1,-2.4]	[-6.7,-1.6]	[-5.1,-2.2]	[-5.1,-2.1]	[-5.1,-2.1]	[-6.1,-2.4]	[-4.9,-0.8]
N. schools	5164	5164	5164	5164	5164	5164	5164	5164	5164	5164
R2	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.8	0.8	0.6
Including state-managed schools that participated of the managerial practices intervention										
	Math					**Portuguese**				
	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D	Triple D
	(4)	(5)	(6)	(7)	(8)	(4)	(5)	(6)	(7)	(8)
H1N1	-4.35***	-4.25***	-4.38***	-4.29***	-4.25***	-3.49***	-3.40***	-3.51***	-3.41***	-2.73**
	(0.86)	(0.92)	(0.87)	(0.93)	(1.30)	(0.75)	(0.80)	(0.75)	(0.81)	(1.13)
p-value	0.000	0.000	0.000	0.000	0.002	0	0.000	0.000	0.000	0.018
95% CI	[-6.1,-2.6]	[-6.1,-2.4]	[-6.1,-2.7]	[-6.1,-2.4]	[-6.8,-1.7]	[-5.0,-2.0]	[-5.0,-1.8]	[-5.0,-2.0]	[-5.0,-1.8]	[-5.0,-0.5]
N. schools	5329	5329	5329	5329	5329	5329	5329	5329	5329	5329
R2	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.8	0.8	0.6
(A) Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(B) Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(C) % teachers in both networks	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes
(D) % teachers management	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
(E) School FE	No	No	No	No	Yes	No	No	No	No	Yes

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the managerial practices intervention. (E) schools' fixed effects.

Table A.16: Impact of the school shutdowns on students' learning (with and without controls), fifth-grade (2007-2009)

	Math						Portuguese					
	DiD	DiD	DiD	DiD	Triple DiD	Triple DiD	DiD	DiD	DiD	DiD	Triple DiD	Triple DiD
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
	No controls		With controls				No controls		With controls			
H1N1	*-3.33***	-3.26**	-3.27**	-2.75**	-2.89*	-4.18***	*0.02*	-0.76	-0.78	-0.54	-1.91*	-2.87***
	(0.94)	(1.18)	(1.22)	(1.23)	(1.49)	(1.28)	(0.81)	(0.86)	(0.91)	(0.95)	(1.13)	(1.05)
p-value	0.01	0.04	0.03	0.04	0.05	0.00	0.98	0.39	0.42	0.61	0.09	0.01
95% CI	[-5.2,-1.5]	[-5.6,-0.9]	[-5.7,-0.9]	[-5.2,-0.3]	[-5.8,0.1]	[-6.7,-1.6]	[-1.6,1.6]	[-2.5,0.9]	[-2.6,1.0]	[-2.4,1.3]	[-4.2,0.3]	[-4.9,-0.8]
N. schools	6416	3912	3912	3912	7329	5164	6416	3912	3912	3912	7329	5164
R2	0.6	0.8	0.8	0.8	0.6	0.7	0.5	0.8	0.8	0.8	0.5	0.6
(A) Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(B) Controls	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	No	Yes
(C) % teachers in both networks	No	No	Yes	Yes	No	Yes	No	No	Yes	Yes	No	Yes
(D) % teachers management	No	No	No	Yes	No	Yes	No	No	No	Yes	No	Yes
(E) School FE	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes

Notes: The authors' estimate is based on data from *Prova Brasil*, Census of Education, and IBGE. Wild bootstrap p-values. 95% CI in brackets. ***, **, and * indicate significance at the 1, 5, and 10% critical levels. Standard error in parenthesis and clustered at the municipality level. All regressions are weighted by fifth-grade enrollment at the school level. Math performance on a scale from 0 to 350 (SAEB scale). Portuguese performance on a scale from 0 to 325 (SAEB scale). The Columns DiD show the estimates differences-in-differences as described in subsection 1.5.1 and Columns Triple D show the estimates for triple difference-in-differences as described in subsection 1.5.2. The sample of municipalities and schools included in the analysis are detailed in Table 1.2. (A) municipal fixed effects. (B) students', teachers', schools' and principals' characteristics. (C) the percentage of teachers that work in a state and a locally-managed school simultaneously. (D) the percentage of teachers working in a state-managed school that implemented the managerial practices intervention. (E) schools' fixed effects. All triple DiD estimates exclude state-managed schools in which the state government implemented the managerial practices intervention.

Table A.17: Average number of state and locally-managed schools included in the triple DiD (2007-2009)

Triple Difference-in-Differences

Only municipalities in which there is at least one state-managed school and one locally-managed school

10 municipalities where the local authorities extended the winter break (G = 0)					*112 municipalities where the local authorities did not extend the winter break (G = 1)*				
	mean	sd	min	max		mean	sd	min	max
Schools					**Schools**				
Locally-managed	84.6	155.8	21	527	Locally-managed	19.7	17.9	1	101
State-managed	90	190.9	5	627	State-managed	9.5	13.3	1	96

Source: Census of Education, 2007-2009.

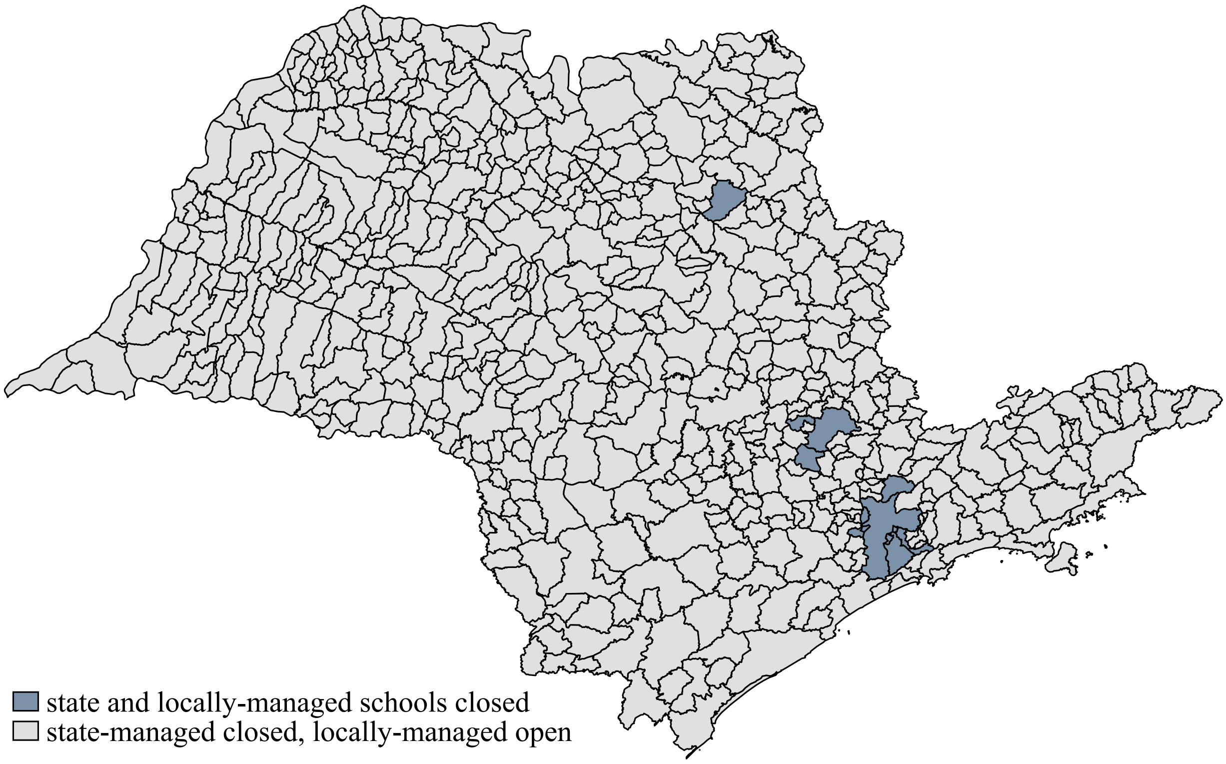
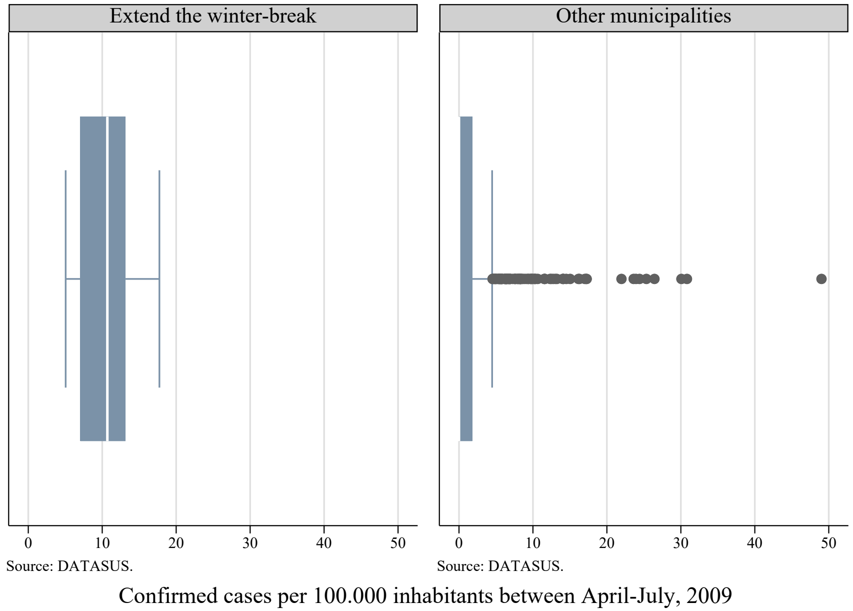
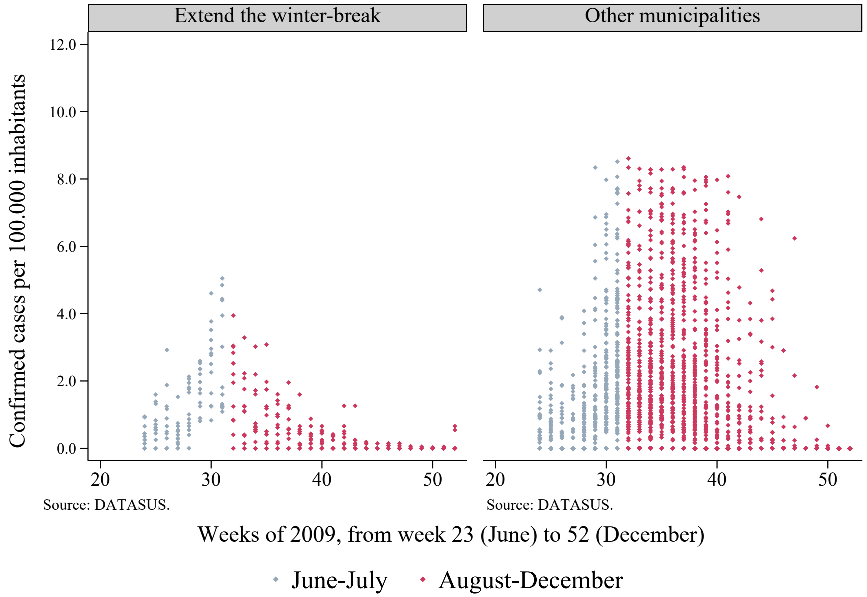


Figure A.1: School shutdown policy, São Paulo (2009)

Source: Authors' analysis based on data from Local newspapers.



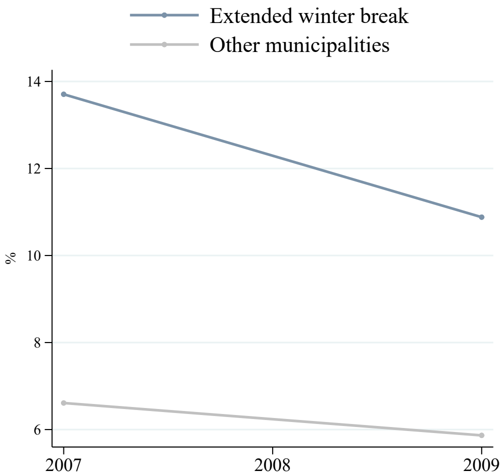
(a) Accumulated, April-July



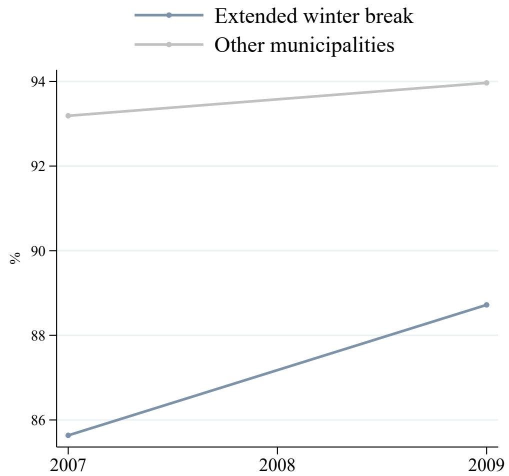
(b) Weekly, April-July

Source: The number of confirmed H1N1 cases per week are from https://datasus.saude.gov.br/informacoes-de-saude-tabnet/. (a) The average number of confirmed H1N1 cases between April and July 2009. (b) The weekly number of cases between April and December 2009.

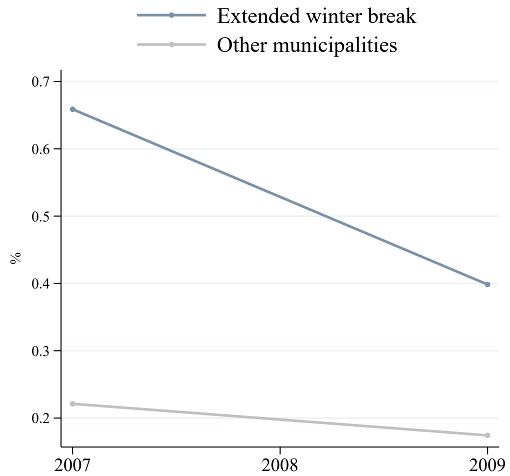
Figure A.2: H1N1 confirmed cases per 100.000 inhabitants, São Paulo (2009)



(a) Retention



(b) Grade-promotion

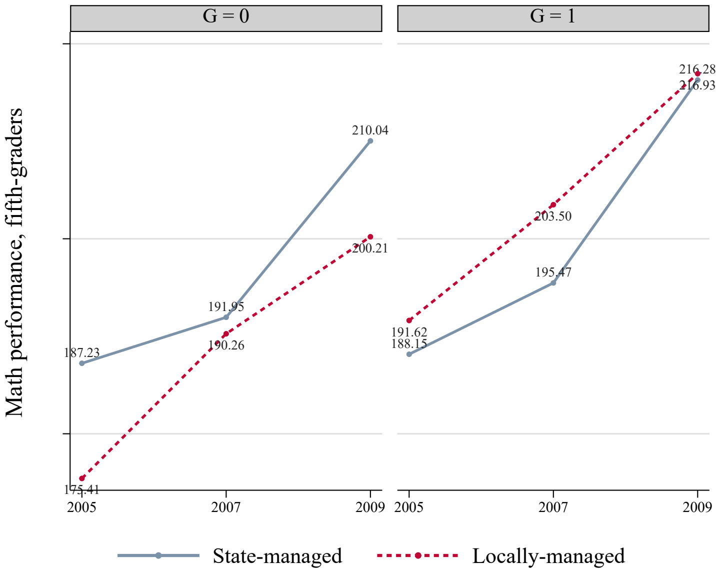


(c) Dropout

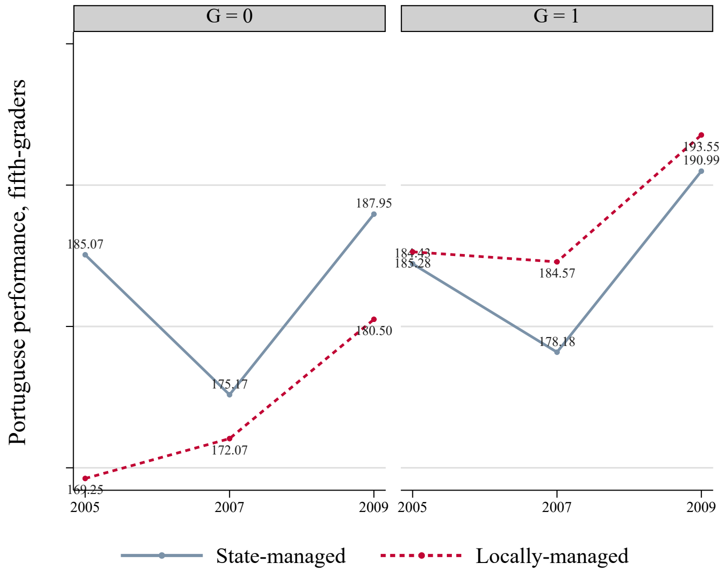
Figure A.3: Retention, dropout and grade-promotion, fifth-grade (2007-2009)

Note: Retention, grade-promotion, and dropout rates range from o to 100. The indicators for treatment and comparison groups are the average for the sample of locally-managed schools in São Paulo. Source: Census of Education/INEP.

Figure A.4: Students' proficiency in state and locally-managed schools, fifth-grade (2005-2009)



(a) Math



(b) Portuguese

Note: The proficiency $G = 1$ and $G = 0$ is the average of the proficiency scores of state and locally-managed schools in São Paulo. The Proficiencies in Portuguese and Math are on the SAEB scale. Source: *Prova Brasil*/INEP.

B Appendix to Chapter 2

Table B.1: Educational indicators, first to fifth grade (2007-2019)

	Age-grade distortion			Retention			Grade-promotion			Dropout		
	Brazil	PE	Recife	Brazil	PE	Recife	Brazil	PE	Recife	Brazil	PE	Recife
2007	25.0	34.3	20.1	12.5	18.3	8.7	84.5	75.5	88.4	3.1	6.2	2.9
2008	19.7	25.8	19.5	11.7	16.8	6.7	85.5	77.8	90.7	2.7	5.4	2.6
2009	20.8	28.5	19.4	10.4	14.0	6.8	87.5	82.0	91.2	2.1	4.0	2.0
2010	20.6	28.0	19.0	9.1	11.8	7.7	89.2	85.1	90.9	1.7	3.1	1.4
2011	19.5	26.2	19.0	7.8	11.1	9.8	90.8	86.4	88.9	1.4	2.5	1.3
2012	18.0	24.4	20.3	7.4	11.6	9.6	91.2	86.0	88.9	1.4	2.4	1.5
2013	16.5	23.1	20.8	6.1	10.3	9.5	92.8	87.9	89.2	1.1	1.8	1.3
2014	14.8	21.7	21.6	6.5	10.6	11.3	92.5	87.7	86.0	1.0	1.7	2.7
2015	13.4	21.0	24.1	6.3	10.3	12.4	92.8	88.2	86.9	0.9	1.5	0.7
2016	12.6	19.6	24.5	6.6	10.5	10.8	92.4	88.1	88.4	1.0	1.5	0.8
2017	12.3	19.2	24.4	5.9	9.2	10.7	93.3	89.7	88.8	0.8	1.2	0.5
2018	11.5	18.1	24.2	5.9	8.0	8.9	93.5	91.1	90.5	0.7	0.9	0.6
2019	10.9	16.7	21.7	5.1	6.6	7.4	94.3	92.8	92.0	0.5	0.7	0.6

Notes: Brazil: average outcomes of locally-managed schools in Brazil. PE: average outcomes of locally-managed schools in the state of Pernambuco. Recife: average outcomes of locally-managed schools in the municipality of Recife. Source: Census of Education/INEP (2007-2019).

Table B.2: Percentage of students in the EMPREL dataset whose proficiency in reading and math is available (2010-2018)

	All			**Regular schools**			**Schools offering Acelera**		
	Second grade	Third grade	Fifth grade	Second grade	Third grade	Fifth grade	Second grade	Third grade	Fifth grade
2010		55.3	59.8		55.8	60.5		49.5	51.0
2011		56.9	70.9		57.7	72.0		54.4	67.9
2012		61.6	64.4		63.3	65.5		58.4	62.5
2013		22.7	56.1		24.4	57.6		18.8	53.0
2014		70.4	70.6		70.9	71.2		68.1	68.6
2015		60.3	61.8		61.9	62.7		56.6	59.8
2016	60.0		64.2	59.3		68.5	61.1		59.2
2017	75.5		79.3	77.1		86.0	73.8		73.4
2018	76.1		78.7	78.5		86.2	74.0		73.3
Total	70.7	54.6	67.2	71.0	56.1	68.3	70.4	49.8	65.1

Notes: The Table shows the percentage of primary education students enrolled in the locally-managed schools of the municipality of Recife that I find in the SAEPE dataset (dataset that has students' proficiency in reading and math in the standardized state exam). Source: EMPREL/Recife and SAEPE/Department of Education of the state of Pernambuco.

Table B.3: Balance test between Regular and *Acelera* schools (2010-2018)

Variable	(1) Regular N	(1) Regular Mean/SE	(2) *Acelera* N	(2) *Acelera* Mean/SE	T-test Difference (1)-(2)
Enrollments, first to fifth graders	1390	235.39	506	312.48	-77.09***
		(3.06)		(5.51)	
Age grade distortion year before, third grade - %	1280	24.10	500	31.33	-7.23***
		(0.30)		(0.48)	
Age grade distortion year before, fourh grade - %	1253	25.54	503	32.64	-7.10***
		(0.32)		(0.52)	
Age grade distortion year before, fifth grade - %	1245	26.58	500	32.22	-5.64***
		(0.30)		(0.50)	
Number of classrooms in the school	1152	8.22	326	10.14	-1.93***
		(0.10)		(0.21)	
Average proficiency Math year before, third grade	681	484.17	259	481.30	2.87
		(1.31)		(1.96)	
Average proficiency Portuguese year before, third grade	833	489.81	304	483.68	6.13***
		(1.24)		(1.83)	
Average proficiency Math year before, fifth grade	1061	190.80	485	190.37	0.43***
		(0.51)		(0.65)	
Average proficiency Portuguese year before, fifth grade	1061	182.85	485	183.09	-0.24***
		(0.50)		(0.71)	
Insufficient score Portuguese year before, third grade - %	681	76.19	259	76.72	-0.52
		(0.54)		(0.80)	
Insufficient score Math year before, fifth grade - %	833	52.80	304	55.91	-3.11***
		(0.57)		(0.83)	
Insufficient score Portuguese year before, fifth grade - %	1061	46.02	485	46.14	-0.11***
		(0.50)		(0.67)	
Insufficient score Math year before, fifth grade - %	1061	46.00	485	45.99	0.00***
		(0.51)		(0.70)	
Grade-promotion year before, first to fifth grade - %	1372	90.35	506	86.56	3.79***
		(0.14)		(0.26)	
Repetition rate year before, first to fifth grade - %	1372	8.34	506	11.96	-3.62***
		(0.13)		(0.25)	
Dropout rate year before, first to fifth grade - %	1372	1.31	506	1.48	-0.17***
		(0.05)		(0.10)	
Library - %	1152	46.96	326	69.33	-22.36***
		(1.47)		(2.56)	
Broadband internet - %	1120	92.41	310	92.90	-0.49
		(0.79)		(1.46)	
Computer Lab - %	1152	75.87	326	81.29	-5.42***
		(1.26)		(2.16)	
Science Lab - %	1152	2.78	326	3.07	-0.29
		(0.48)		(0.96)	
Number of classrooms available morning year before	1229	1.95	406	2.34	-0.39***
		(0.06)		(0.11)	
Number of classrooms available afternoon year before	1227	2.16	400	2.39	-0.23**
		(0.07)		(0.11)	

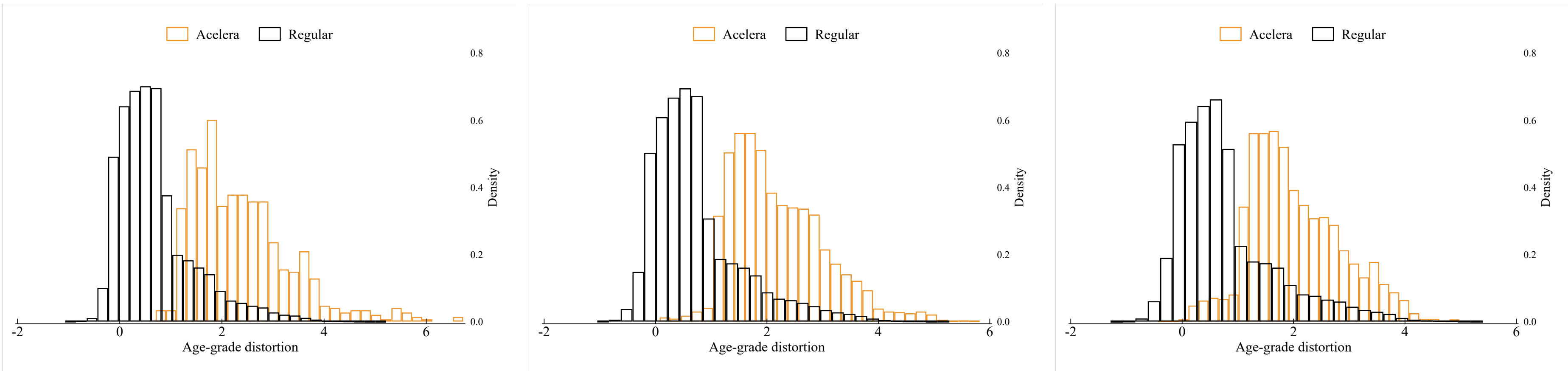
Notes: The value displayed for t-tests is the differences in the means across the groups. Year-fixed effects are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical levels. The Table shows data for schools that only offer Regular Education and schools that offer *Acelera* and Regular Education. The variables labeled with *year before* refer to the educational indicators one year before the school is selected (or not) to implement *Acelera*. Source: EMPREL, SAEPE, and Census of Education.

Table B.4: Balance test between *Acelera* participants that have and that do not have proficiency in fifth grade available (2010-2018)

	1		2		t-test
	Did not participate SAEPE 5th grade		Participated SAEPE 5th grade		Difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Age-grade distortion, 4th grade	2650	2.30	431	2.09	0.21***
		[0.02]		[0.04]	
Approval rate, 4th grade	2515	88.71	479	86.22	2.49
		[0.63]		[1.58]	
Repetition rate, 4th grade	2515	8.63	479	13.57	-4.94
		[0.56]		[1.57]	
Performance in Portuguese, 3rd grade	1318	459.91	279	448.09	11.82***
		[2.50]		[5.40]	
Performance in Math, 3rd grade	1168	460.84	256	457.25	3.59*
		[2.54]		[5.53]	
If approved in 4th grade, jumped to 6th grade	1086	50.55	478	0.00	50.55***
		[1.52]		[0.00]	

Notes: The Table only includes data for children that joined *Acelera* in fourth-grade. I compare the students that have fifth-grade performance in reading and math with the ones that do not. Source: EMPREL and SAEPE.

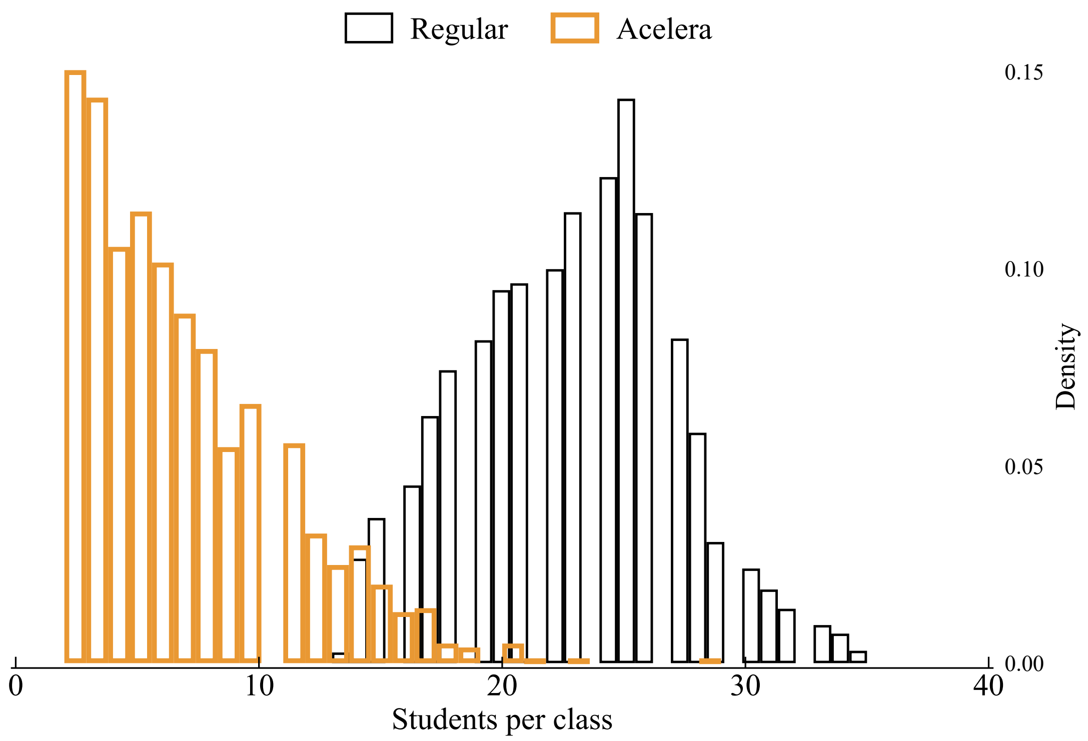
Figure B.1: Distribution of students with at least one year of age-grade distortion, third to fifth grade (2010-2018)



(a) Third-grade (b) Forth-grade (c) Fifth-grade

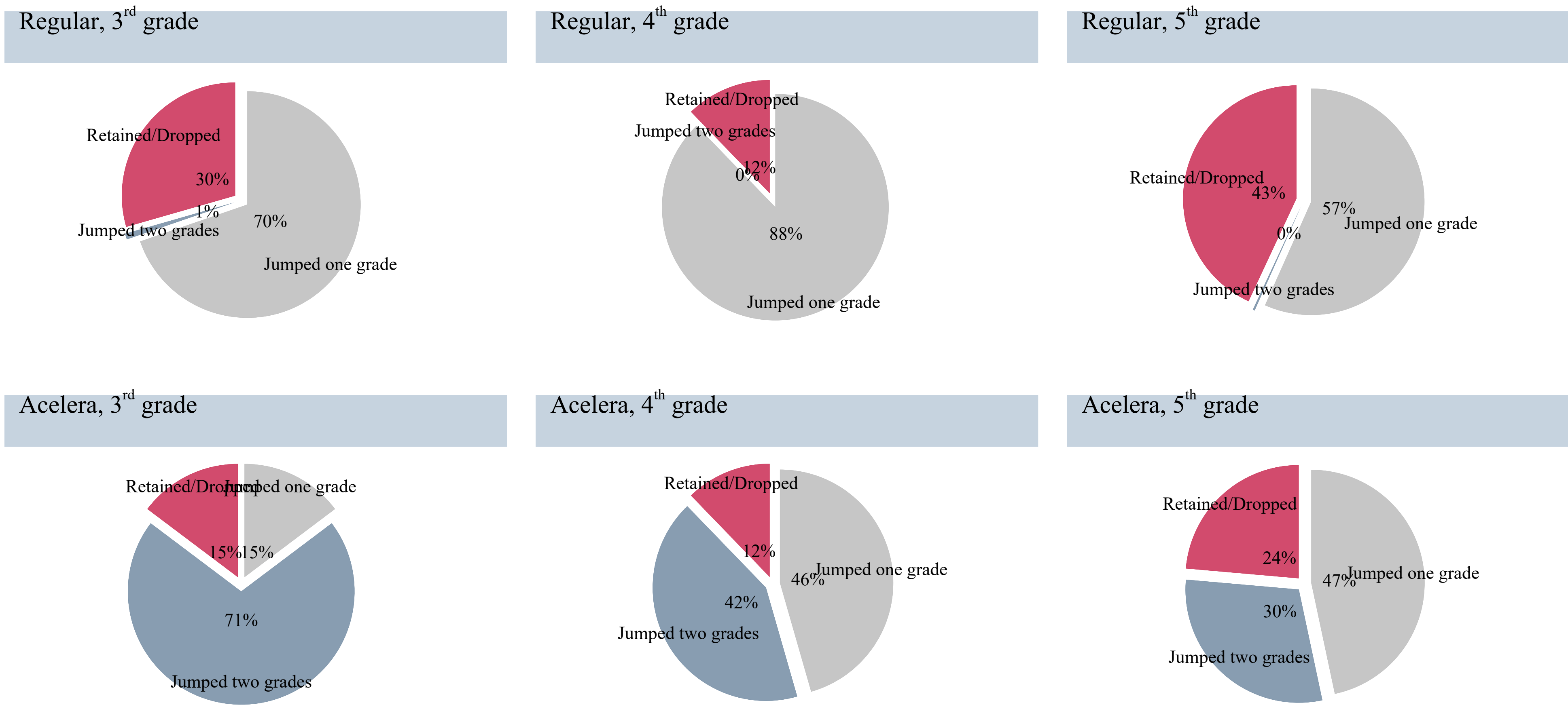
Note: The Figures show the enrollment of students with at least one year of age-grade distortion in Regular Education and in *Acelera*. Source: EMPREL.

Figure B.2: Distribution of students per class, third to fifth grade (2010-2018)



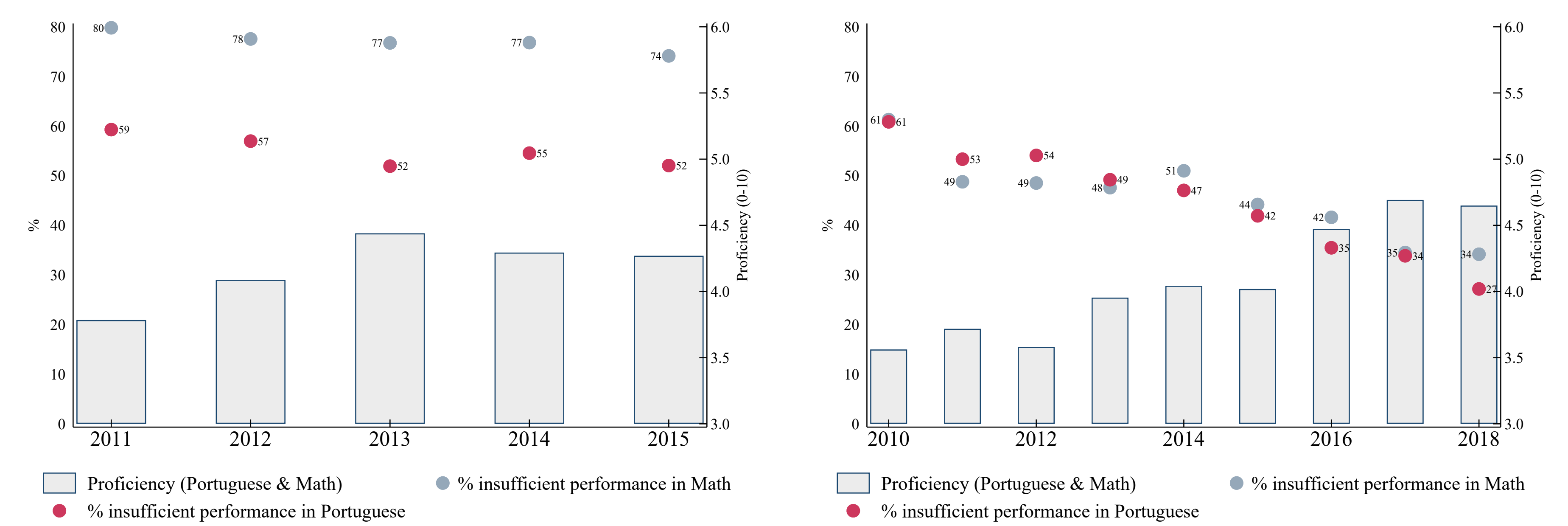
Note: The Figure shows the distribution of students per class in Regular Education and in *Acelera*. Data from third, fourth, and fifth grades. Source: EMPREL.

Figure B.3: Retention, dropout and grade-promotion of Regular and *Acelera* students, third to fifth grade (2010-2014)



Note: Dropout, grade promotion, and retention rates of students with at least one year of age-grade distortion. Source: EMPREL.

Figure B.4: Performance in reading and math, third and fifth grade (2010-2018)

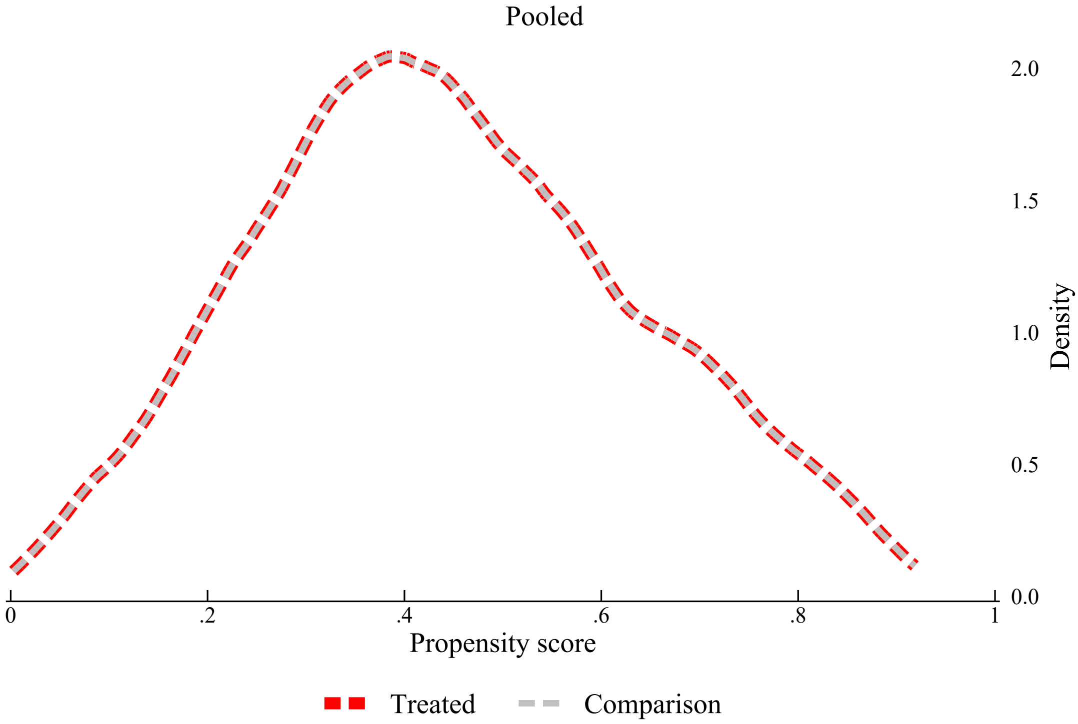


(a) Third-grade

(b) Fifth-grade

Note: The performance of third-graders is only available until 2015. Source: EMPREL.

Figure B.5: Propensity score matching on the sample of schools offering *Acelera*, first to fifth grade (2010-2014)



Note: Authors' estimate. I match students with at least one year of age-grade distortion who are enrolled in regular education with *Acelera* students. I use as controls: the age-grade distortion (in years), school of enrollment, grade, sex of the student, the difference in years between the oldest and the youngest of the classroom the student is enrolled in, and whether or not the student already participated of *Se Liga*. The sample contains only data from schools that offer *Acelera* in at least one year between 2010 and 2014. Source: EMPREL and SAEPE.

C Appendix to Chapter 3

Table C.1: Price deflator by geographical area (2008)

Geographical context	**Price deflator**
Metropolitan urban area of Belém	0.93
Urban North excluding metropolitan urban areas	0.94
Rural North	0.89
Metropolitan urban area of Fortaleza	0.90
Metropolitan urban area of Recife	0.89
Metropolitan urban area of Salvador	0.98
Urban Northeast excluding metropolitan urban areas	0.90
Rural Northeast	0.86
Metropolitan urban area of Belo Horizonte	1.03
Metropolitan urban area of Rio de Janeiro	0.96
Metropolitan urban area of São Paulo	1.00
Urban Southeast excluding metropolitan urban areas	0.96
Rural Southeast	0.92
Metropolitan urban area of Curitiba	0.95
Metropolitan urban area of Porto Alegre	1.00
Urban South excluding metropolitan urban areas	0.95
Rural South	0.82
Brasília	1.02
Urban Midwest excluding Brasília	1.01
Rural Midwest excluding Brasília	0.95

Source: Brazilian Expenditure Survey (POF) disclosed by Institute of Geography and Research (2008-2009).

Table C.2: Descriptive statistics of the inputs and outputs according to municipality size (2015-2019)

		Population size					
		Up to 5k	Between 5k and 50k	Between 50k and 100k	Between 100k and 500 k	More than 500k	Total, Brazil
2019 IDEB 1st to 5th grade	mean	5.91	5.59	5.72	5.91	5.73	5.68
	sd	0.94	1.01	0.94	0.80	0.71	0.99
	min	3.30	2.30	3.60	3.60	4.60	2.30
	max	8.50	9.40	8.90	8.40	7.40	9.40
Expenditure per student in 2019	mean	11,847	8,404	8,028	8,441	10,064	9,176
	sd	4,557	2,459	2,222	2,292	2,882	3,378
	min	5,157	4,811	4,809	4,920	5,859	4,809
	max	75,711	23,840	17,210	18,877	20,873	75,711
Expenditure per student in 2018	mean	11,587	8,069	7,704	8,083	9,745	8,858
	sd	4,591	2,465	2,541	2,197	2,934	3,416
	min	4,749	4,581	4,582	4,655	5,508	4,581
	max	72,653	21,256	31,702	18,115	20,030	72,653
Expenditure per student in 2017	mean	10,478	7,724	7,459	7,810	9,415	8,348
	sd	3,868	2,049	1,862	1,893	2,858	2,817
	min	4,455	4,361	4,370	4,631	5,249	4,361
	max	56,255	22,325	12,241	16,785	20,640	56,255
Expenditure per student in 2016	mean	10,729	7,794	7,347	7,707	9,964	8,447
	sd	4,109	2,087	1,917	2,079	3,339	2,965
	min	5,181	4,632	4,634	4,621	5,038	4,621
	max	51,347	24,268	14,277	14,855	21,112	51,347
Expenditure per student in 2015	mean	12,143	8,435	7,908	8,362	10,435	9,254
	sd	6,291	2,957	2,440	2,394	3,309	4,229
	min	5,081	4,833	4,974	4,943	5,512	4,833
	max	85,104	46,084	19,737	18,825	20,842	85,104
2019 enrollments 1st to 5th grade	mean	210	1,058	3,875	9,715	42,854	1,829
	sd	100	813	1,577	5,160	48,188	6,302
	min	11	56	913	1,683	9,545	11
	max	922	7,007	9,765	28,515	263,392	263,392
% of students' mothers with high school	mean	50.1	44.6	46.1	52.1	51.7	46.3
	sd	17.3	13.8	10.2	9.0	8.3	14.4
	min	0.0	0.0	20.8	22.0	38.8	0.0
	max	100.0	100.0	70.3	82.4	69.8	100.0

Source: The Education Development Index (IDEB) comes from the Education Assessment System (SAEB) disclosed by the National Institute of Educational Studies and Research (INEP). IDEB of schools managed by local authorities. The percentage of students whose mothers finish high school comes from the socioeconomic questionnaire of *Prova Brasil* applied by the INEP. The per-pupil expenditure is shown in 2020 BRL and comes from the Information System on Expenditures in Education (SIOPE). It refers to the expenditures made by the municipalities in their locally-managed schools. The expenditure per student considers the difference in the cost of living in metropolitan, urban and rural areas according to the deflators shown in Table C.1.

Table C.3: Sample-size of the Brazilian municipalities included in the DEA (2019)

	Number of municipalities				**Enrollments, first to fifth grade**			
	Without	**Excluding outliers**			**Without**	**Excluding outliers**		
	exclusions	**Super-efficiency thresholds**			**exclusions**	**Super-efficiency thresholds**		
Number of inhabitants		**1.3**	**1.2**	**1.1**		**1.3**	**1.2**	**1.1**
Up to 5k	960	953	944	935	219,830	218,011	216,186	213,915
Between 5k and 50k	3,220	3,217	3,211	3,200	3,414,887	3,413,028	3,409,427	3,397,576
Between 50k and 100k	330	328	326	313	1,252,764	1,247,760	1,244,163	1,185,128
Between 100k and 500 k	253	250	250	243	2,445,813	2,428,771	2,428,771	2,381,263
More than 500k	44	40	40	37	1,955,619	1,844,189	1,844,189	1,774,996
Total, Brazil	4,807	4,788	4,771	4,728	9,288,913	9,151,759	9,142,736	8,952,878

Source: The number of municipalities comes from the Brazilian Institute of Geography and Research (IBGE). The enrollments are from the Census of Education (2019) and include only schools managed by local authorities. The columns "without exclusions" show the number of municipalities before the exclusion of outliers. The super-efficiency columns indicate the number of DMUs in the models that exclude outliers. The thresholds used are 1.5, 1.3, and 1.1. See Section 3.4.

Table C.4: Number of benchmark DMUs by state (2019)

Brazilian region	State	N. municipalities	Group Frontier Benchmarks	%	Meta Frontier Benchmarks	%
North	AC	19	4	21	3	16
North	PA	86	18	21	8	9
North	AM	48	9	19	5	10
Northeast	MA	167	23	14	14	8
Northeast	AL	77	10	13	6	8
Southeast	ES	65	8	12	4	6
Northeast	CE	177	19	11	10	6
Northeast	PI	189	17	9	8	4
Southeast	MG	740	65	9	32	4
Northeast	PB	197	16	8	7	4
Southeast	RJ	78	6	8	1	1
Northeast	PE	174	13	7	5	3
South	PR	381	23	6	12	3
Northeast	BA	382	18	5	10	3
Northeast	SE	67	3	4	1	1
Midwest	GO	225	10	4	2	1
Southeast	SP	568	25	4	4	1
North	RO	50	2	4	2	4
Northeast	RN	131	5	4	2	2
Midwest	MT	114	4	4	1	1
South	SC	259	7	3	4	2
South	RS	319	6	2	2	1
North	TO	121	2	2	0	0
North	AP	13	0	0	0	0
North	RR	8	0	0	0	0
Midwest	MS	73	0	0	0	0
Total	Brazil	4,728	313	193	143	97

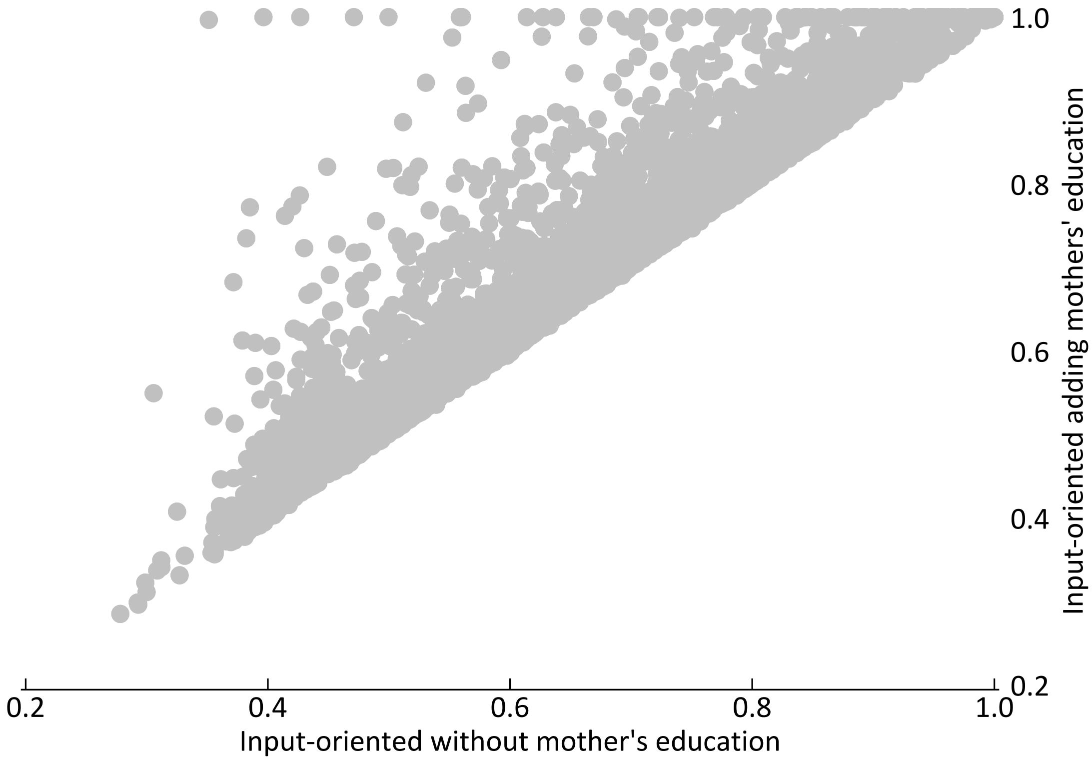
Source: The authors' estimate is based on *Prova Brasil*, IBGE, and SIOPE. Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1.

Table C.5: Input-oriented efficiency scores (2013-2019)

	Group Frontier		Meta Frontier	
Number of inhabitants	2013	2019	2013	2019
Up to 5k	0.73	0.72	0.60	0.60
Between 5k and 50k	0.73	0.74	0.73	0.74
Between 50k and 100k	0.83	0.83	0.73	0.74
Between 100k and 500 k	0.79	0.80	0.69	0.72
More than 500k	0.78	0.83	0.54	0.56

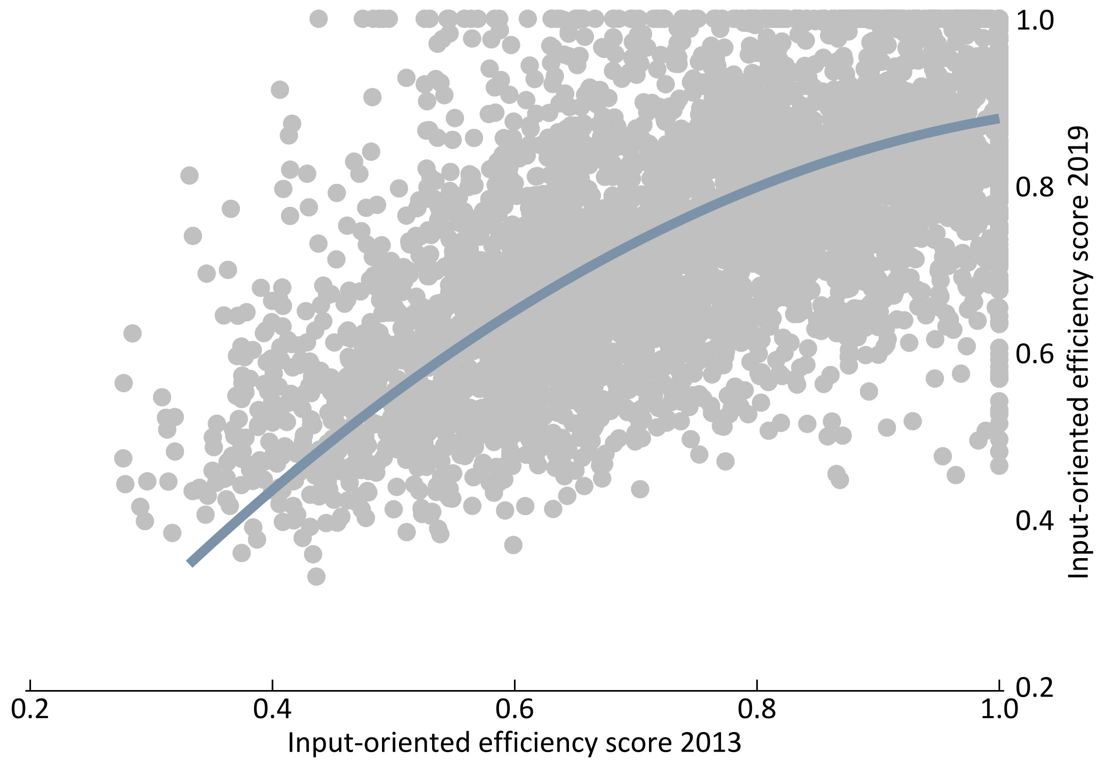
Source: The authors' estimate is based on 2019 *Prova Brasil*/INEP, and 2015-2019 SIOPE. Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1.

Figure C.1: Distribution of input-oriented efficiency scores, with and without non-discretionary inputs (2019)



Note: The authors' estimate is based on *Prova Brasil*, IBGE, and SIOPE. Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1. The x-axis shows the input-oriented score of a model in which the output is the primary education IDEB and the inputs are the expenditure per student in 2015, 2016, 2017, 2018, and 2019. The y-axis is similar but the analysis includes the percentage of students whose mothers finish high school as a non-discretionary input. I observe that the model that does not include mothers' education significantly underestimates the efficiency of the municipalities.

Figure C.2: Distribution of input-oriented efficiency scores (2013-2019)



Note: The authors' estimate is based on*Prova Brasil*/INEP, and SIOPE. Municipal governments. Estimates on the sample of DMUs that have super-efficiency scores lower than 1.1. The x-axis shows the input-oriented score using the 2013 primary education IDEB. The y-axis shows the input-oriented score using the 2019 primary education IDEB. Overall, the higher the efficiency in 2013, the higher the efficiency in 2019.