FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

LEONARDO MAKOTO KAWAHARA

UNDERSTANDING THE IMPACTS OF A FULL-TIME SCHOOLING PROGRAM

São Paulo

2019

LEONARDO MAKOTO KAWAHARA

UNDERSTANDING THE IMPACTS OF A FULL-TIME SCHOOLING PROGRAM

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Orientador: Vladimir Pinheiro Ponczek.

.

São Paulo

2019

Kawahara, Leonardo Makoto.
Understanding the impacts of a full-time schooling program / Leonardo Makoto Kawahara. - 2019.
91 f.

Orientador: Vladimir Pinheiro Ponczek.
Dissertação (mestrado CMEE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Escolas públicas - Brasil. 2. Educação integral. 3. Avaliação educacional - Brasil. I. Ponczek, Vladimir Pinheiro. II. Dissertação (mestrado CMEE) – Escola de Economia de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 37.018.591(81)

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O
Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

LEONARDO MAKOTO KAWAHARA

UNDERSTANDING THE IMPACTS OF A FULL-TIME SCHOOLING PROGRAM

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Data de aprovação: __/__/____

Banca examinadora:

Prof. Dr. Vladimir Pinheiro Ponczek
FGV-EESP (Orientador)

Prof. Dr. André Portela Fernandes de Souza
FGV-EESP

Prof. Dr. Naercio Aquino Menezes Filho
Insper

Dedico este trabalho ao meus pais, Paulo e Suely,
e à minha irmã, Paula,
por estarem ao meu lado em todos os momentos da minha vida
e por acreditarem em mim.

Aos meus amigos da minha turma de pós-graduação,
em especial Gabriel, Bruna, Túlio, Nicolas
e João Pedro, pelos desafios que juntos superamos.

Aos meus amigos da salinha.

Aos meus amigos Janaína, Pedro, Rodrigo, Adriano e Arthur,
por estarem do meu lado nos momentos de maior necessidade.

Aos meus primos, Marina, Melina e Samuel,
pelo apoio.

À Erica, pelo companheirismo.

Agradecimentos

Ao meu orientador, Vladimir Ponczek, pelo esforço e pela orientação.

À Helena Lima e à Thereza Barreto, pelo auxílio, informações e disponibilidade.

Ao corpo docente da Escola Economia de São Paulo da Fundação Getúlio Vargas pelas aulas e pelo conhecimento fornecido.

"O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001."

"This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001"

Resumo

Este trabalho estima os impactos do programa de ensino em tempo integral, uma intervenção em escolas públicas brasileiras de ensino fundamental e médio, em notas médias da Prova Brasil e do ENEM, e em indicadores de fluxo, como taxas de aprovação, reprovação, abandono e distorção série-idade. Para atingir tal objetivo, fizemos o uso da metodologia de diferenças-em-diferenças com dados do Censo Escolar, ENEM, Prova Brasil, e indicadores de fluxo. Os resultados indicam que o programa tem impactos significantes em todos as variáveis. Em média, o programa aumenta as notas das escolas, aumenta a taxa de aprovação e diminui as taxas de reprovação, desistência e distorção série-idade. Testamos para outros efeitos do programa, como impacto na composição da escola e em escolas não tratadas na mesma cidade. Os resultados sugerem que existe um efeito seleção em ação, mas que não é o maior responsável pelo impacto do programa. Ainda, buscando entender os canais de transmissão do programa, usamos variações na implementação do programa para estimar quais características parecem ser mais relevantes para o impacto positivo. Os resultados indicam que, em escolas que implementaram processos de seleção e remoção de professores e diretores, o programa teve maior impacto. Essas descobertas trazem maior compreensão em como programas de ensino em tempo integral impactam as variáveis educacionais.

Palavras-chave: Ensino em tempo integral, aumento da jornada escolar, qualidade do ensino.

Abstract

This work estimates the impact of the full-time schooling program, an intervention in Brazilian public middle and high schools, on test score such as *Prova Brasil* and ENEM test scores and on flow indicators, such as promotion, failure, drop-out and grade-age distortion. The results indicate that the program has significant impacts in all the outcomes. In average, the program increases the test scores of schools, increases promotion rates, decreases failure, drop-out, and grade-age distortion rates. We test for other effects of the program, like the impact on school composition and on other schools in the city. The findings suggest that there is a selection effect in place, but that they are not the main responsible for the impact of the program. Furthermore, to understand the channels of transmission of the program, we use some variation in the implementation of the program to estimate which characteristics seem to be more relevant for the positive impact. The results indicate that, in schools which had implemented process of selection and removal of teachers and principals, the impact of the program was larger. These findings bring comprehension in how programs of full-time schooling may impact educational outcomes.

Keywords: Full-time schooling, quality of education, extension of school time.

List of Figures

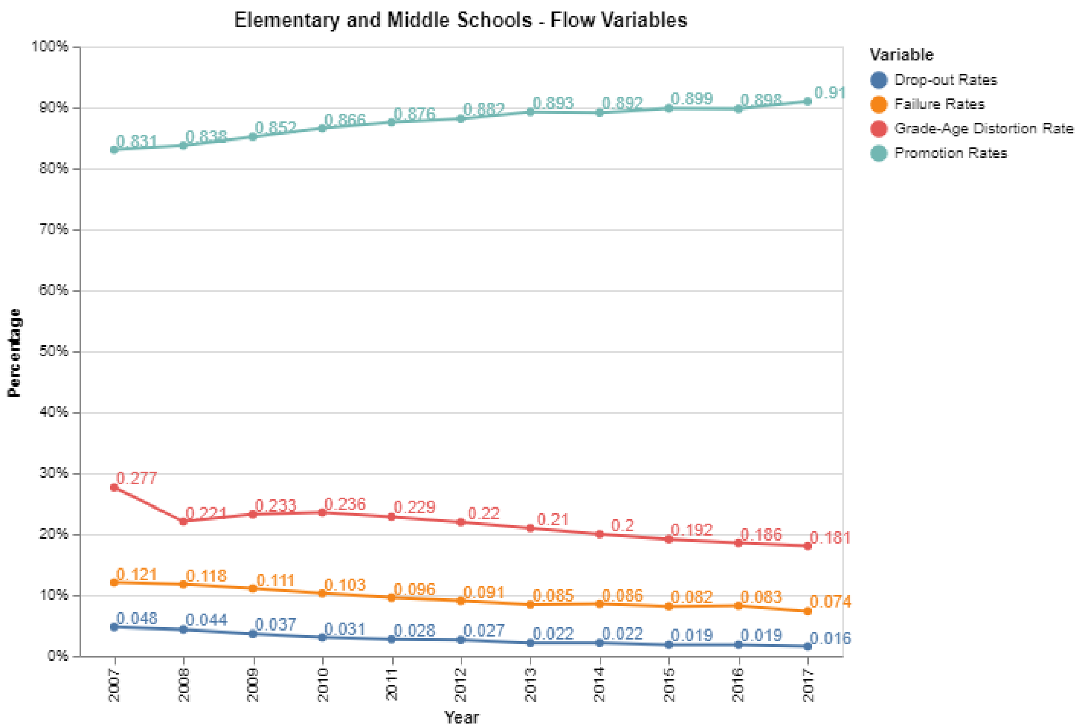
List of Tables

Contents

1 Introduction

In the past 15 years, the situation of education in Brazil has been characterized by the overcoming of some challenges, but also the failure to address others. According to PNAD[1], the percentage of children from 6 to 14 years old enrolled in primary school and middle school increased from 91.5% in 2004 to 97.8% in 2015 (see graph 3). Similarly, according to INEP[2], in primary education, from 2007 to 2017, promotion rates increased, and failure, drop-out and grade-age distortion rates. These can be seen in the graph 1. The trends of these variables indicate that, for primary education in Brazil, challenges like low enrollment and high drop-out rates are or are in the process of being dealt with.

Figure 1 – Elementary and middle schools flow variables



Sources: INEP

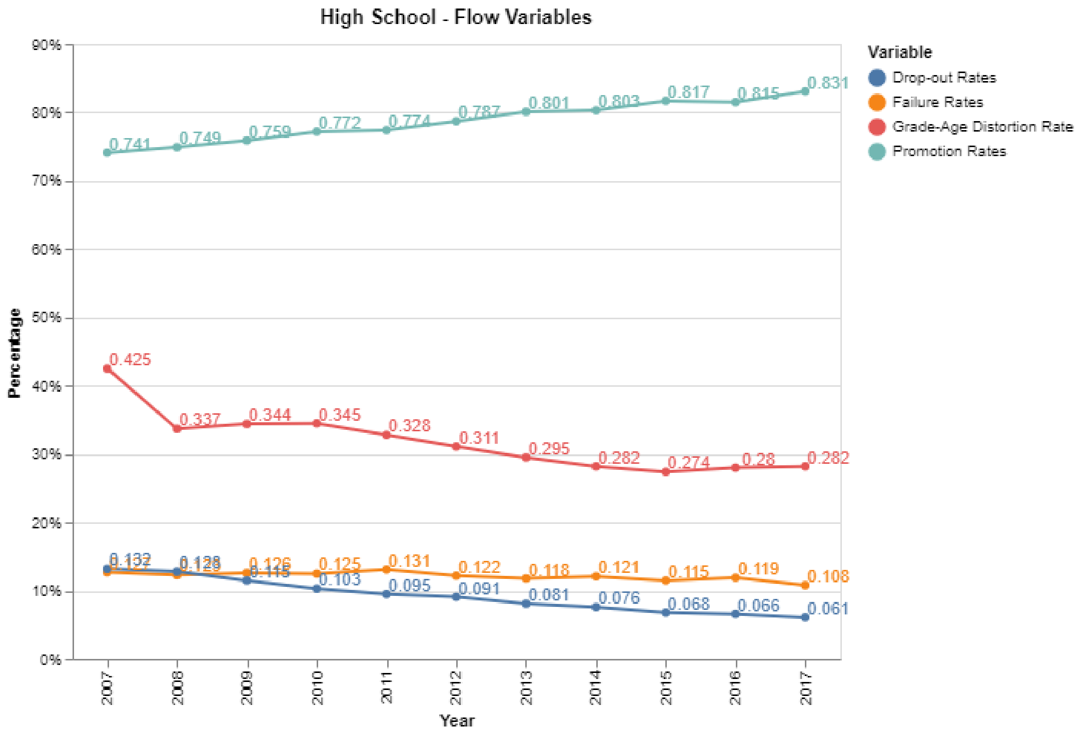
For secondary education, the scenario is to some degree similar. Looking at the analogous variables, we have that the percentage of the population from 15 to 17 that either attended or concluded high school increased from 51.1% in 2004 to 70.1% in 2017 (see graph 4); from 2008 to 2017, the promotion rates increased, failure, drop-out, and age-grade distortion rates decreased. These can be seen in the graph 2. The trends of these

[1] *Pesquisa Nacional Adicional de Domicílios (PNAD)*, National Households Survey, in portuguese

[2] *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*, an agency related to Ministry of Education which evaluates education in Brazil

variables are slower when compared to the ones for primary education, but, assuming that all relevant variables for these trends are kept constant, problems such as high drop-out and grade-age distortion rates will eventually be solved.

Figure 2 – High schools flow variables



Sources: INEP

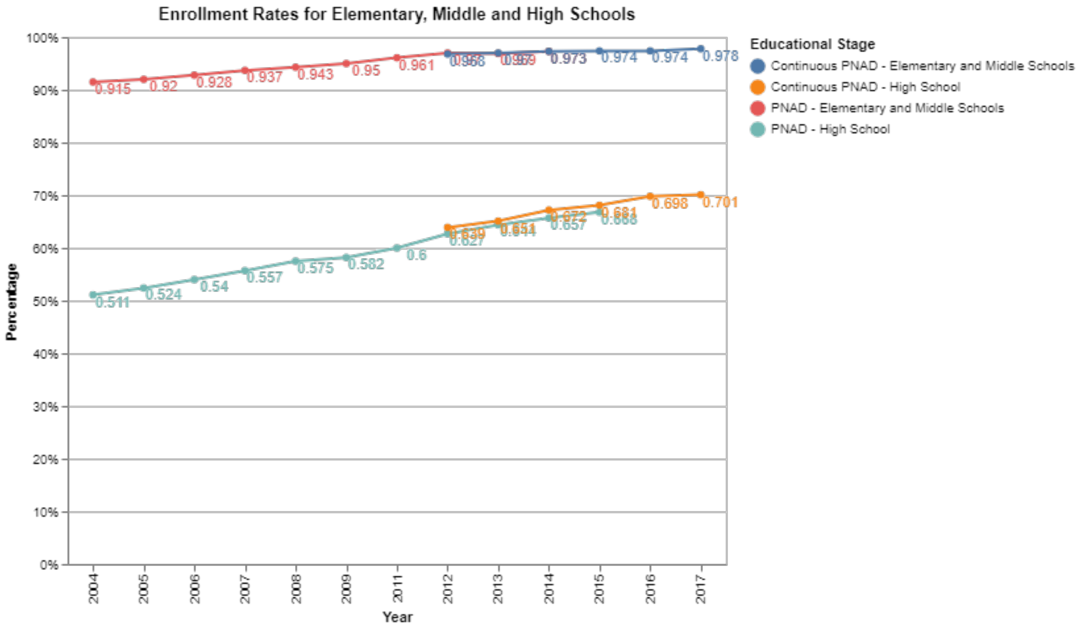
However, when it comes to quality, the scenario is very different. Graphs 5 and 6 show the evolution of average proficiency scores in SAEB[3] for the 5th grade of primary school, 9th grade of middle school, and 3rd grade of high school, for subjects of Portuguese language and mathematics. The scores range from 0 to 350, for the 5th grade; 200 to 425 for the 9th grade; and from 225 to 475 for the 3rd grade. For the 9th grade and for the 3rd grade, in both subjects, the trends of average scores either stayed flat or went down in 22 years of evaluation. For the 5th grade, in both subjects, the trends went up 15%, more or less. Add this to the fact that the first scores are already in a very low level[4]. The quality of education is bad, and it does not seem that it is going to improve: the trends indicate that, at best, maintaining all other things constant, quality of Brazil will stagnate in a low baseline.

Additionally, not only in reference to its past values the quality of education in Brazil does not look good, but also when comparing with other countries. The country's

[3] *Sistema Nacional de Avaliação da Educação Básica* (SAEB) is a set of evaluations of the education in Brazil.

[4] For each subject and grade, the scores are divided in 7 to 10 levels. Each level indicates certain competences and capabilities that the student has. In the first years, the scores correspond to level 3 or level 4.

Figure 3 – Enrollment rates



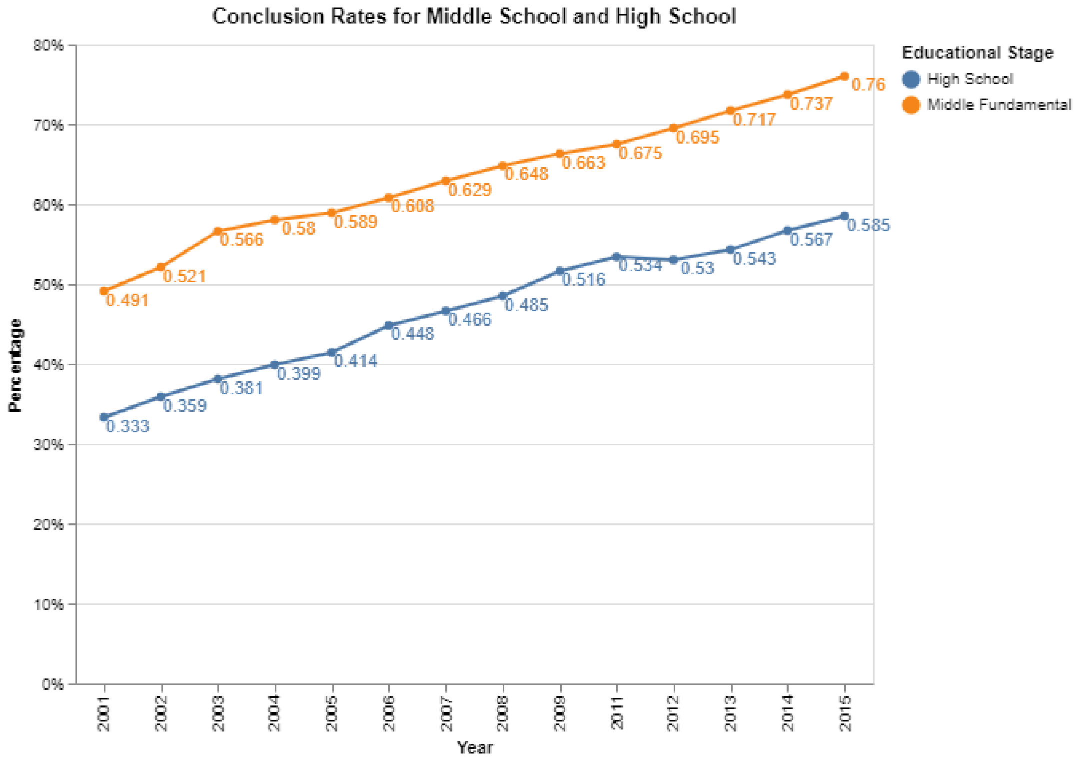
Sources: PNAD/IBGE

performance in PISA[5] is relatively bad. In the 2015 edition, Brazilian students average score in science was 401 when the OECD average was 493, in reading was 407, when the OECD average was 493, and in mathematics was 377, when the OECD average was 490, placing Brazil in the top 10 worst average scores in a set of 72 countries evaluated.

This contrasts with the proportion of GDP spent in education: Brazilian government spends 5.4% of GDP, higher than the average of OCDE (4.8%), and above than other similar countries such as Argentina (4.9%), Chile (4.0%), Colombia (4.2%) and Mexico (4.6%) (OECD (2017))). Furthermore, average skills have improved in comparison to last editions of PISA, however, at their current rate of improvement, they won't reach the average math score of OCDE for 75 years, neither the average language score of OCDE for 260 years (World Bank (2018)).

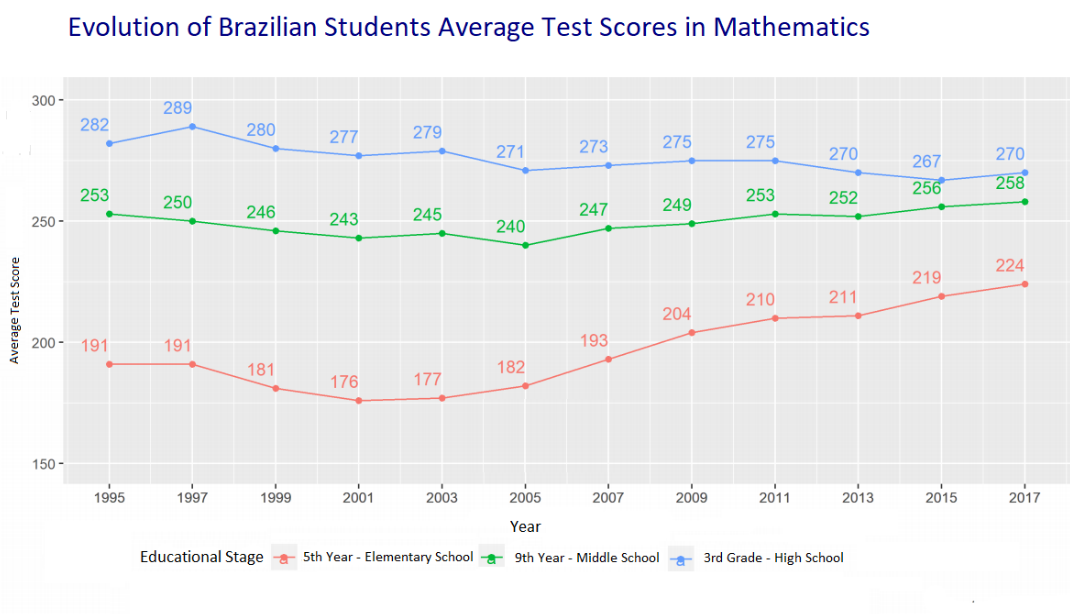
[5] Program for International Students Assessment(PISA) is an assessment of the extent to which 15 years old students of the participating countries have acquired key knowledge and skills that are essential for full participation in modern societies (OECD (2015)).

Figure 4 – Conclusion rates



Sources: PNAD/IBGE

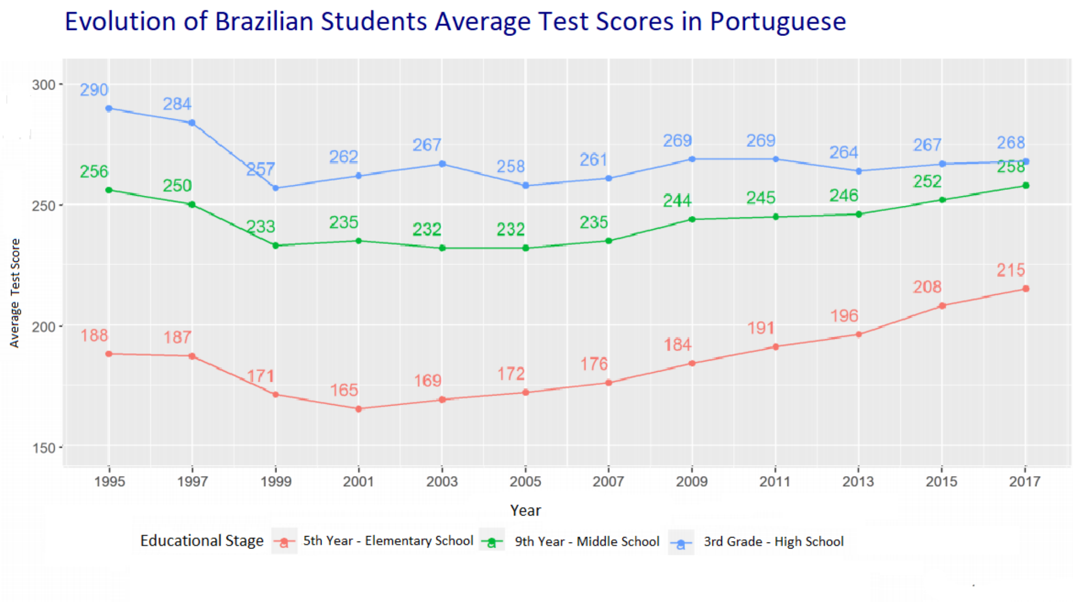
Figure 5 – SAEB mathematics test scores



Sources: DAEB/INEP

These facts portrait a scenario for education in Brazil where challenges such access to schools, enrolment and drop-out have been solved or are in the process of, but the quality issue is still a big pending problem. The Government and society in Brazil are struggling to overcome it. Many initiatives and public policies are being created to improve the overall situation.

Figure 6 – SAEB portuguese test scores



Sources: DAEB/INEP

With the objective of improving quality of education in Brazil, policies involving the "*Ensino em Tempo Integral*[6]", that is, Full-time schooling and extension of school day programs, are starting to gain more notoriety and support. This can be seen in the Goal 6 of the National Plan of Education[7]: "to offer full-time schooling to, at least, 50% of public schools(...)"; in the *Mais Educação* federal program; and in the 2018 secondary school reform, with the program to foster the Implementation of Full-time schooling[8]. Media also has its role. In recent news and debates, *Ensino Integral* is usually seen as a solution for quality education in Brazil.

There are several reasons for the great appeal of full-time schooling in Brazil. First, half-day schooling is the rule: the amount of time of classes required by law is 4h per day. Assuming that time is a critical input for learning, extending to full-day schooling seems a logical step to improve quality, even more when the problem of access seems to have

6 The literal translation would be teaching in integral time. It derives from "Integral Education", a theoretical concept initially used for a type of education whose focus is the development of a integral, whole formation of the human being. One important aspect of the *Ensino Integral* is teaching in integral time, that is full-time schooling.

7 "*Plano Nacional de Educação* (2014)"

8 *Política de Fomento à Implementação de Escolas de Ensino Médio em Tempo Integral*

been taken care of. Second, keeping children in the school afternoon has a great political appeal, since the school has to provide child care and school lunch (Alfaro et al. (2015)). Third, Brazil is going through a major demographic transition. The country is undergoing a unique moment in which the number of new students is decreasing, and, as the total expenditure with education as a proportion of the GDP is kept relatively constant, the expenditure per student will grow. This opens space for more spending per student.

In a more micro level, there is a certain program that is expanding more and more its coverage across Brazil. "The full-time schooling program", as we are going to call it, is a public private initiative that implements new pedagogic and management practices in public schools. In its essence, it is a program of *Ensino Integral*, that tries to approach the education from various aspects, including the extension of school day to full-time schooling (9 hours per day). Hence, full-time schooling for the name of the program. The pedagogic practices consist in principles such as the integrality of education, the youth protagonism, and pedagogy of presence. The management practices consists in the adoption of system of incentives for teachers and employees, adoption of criteria for selection and removal of teachers, adoption of a management instrument derived from the business sector, among others.

In the literature, impacts of full time schooling and extension of school day on outcomes such as test scores are somewhat ambiguous. Theoretically, this is expected. How can extension of school day and more time, can affect the school and the students? More time in school provide more opportunities for longer classes, particularly in core academic subjects; offer teachers extra period for planning and professional development; enable the schools to offer extracurricular activities; and make it possible for schools to allocate time for one-on-one or small group tutoring (Farbman and Kaplan (2005)). In developing countries context, additional hours in school and the consequently provision of school lunches may lead to an increase in enrollment or a decrease in drop out. Also, children would spend more time in a supervised setting, with less exposure to risk behaviours, such as unsafe sex, crime and substance abuse (Alfaro et al. (2015)).

Conversely, more time can have a negative impact. It can cause boredom or school revulsion (Abbiati and Meroni (2016)). If the extra time is used for non-academic subjects, the effect on student learning could be neutral. Similarly, if teachers are not compensated for the additional labor, average teacher effort could decrease in total (Alfaro et al. (2015)). There even the possibility of replacement of home study (Almeida et al. (2016)).

This ambiguity is translated to the empirical works as well. In USA, many studies analyze the impacts of full-day kindergarten versus half day kindergarten. Most of them use the data from the Early Childhood Longitudinal Program, and all of them indicate that children in full-day kindergarten learns more than their half-day counterparts, but this difference fades-out after some years (DeCicca (2007), Votruba-Drzal et al. (2008),

Burkam et al. (2006), Cannon et al. (2006), Zvoch et al. (2008), Robin et al. (2007), Gibbs (2014)). However, still in USA, in other levels of education, the results indicates that more time does not necessarily implies in better outcomes. In James-Burdumy et al. (2005), the impact of after-school program was evaluated , and results indicate that students that attend a after-school learning center were not more likely to have higher academic achievement. In Link and Mulligan (1986), little benefit was expected from a increase in the amount of math and reading instruction. Even further, when evaluating the shortening of the school week, Anderson and Walker (2015) find a positive relationship between the four-day week and performance in reading an mathematics, that is, there is little evidence that moving to four-day week compromises student academic achievement. In other developed countries, we have the same pattern. In a randomized trial in Netherlands, Meyer and Van Klaveren (2013) found that vouchers of participation to a extension of school day program has no impact in math and language tests. In the Quality for Merit Project, a large scale intervention in Southern Italy which provides longer instruction time in low achieving schools, Battistin and Meroni (2016) and Meroni and Abbiati (2016) found that increase in time only led to an increase in math scores.

Focusing mainly in developing countries, Alfaro et al. (2015), in their survey, analyze a diverse set of 19 impact evaluations of school day extension on academic, labor and social outcomes, in Latin America and the Caribbean. Overall, extending the school day tend to show positive results, but they are decidedly mixed, since sometimes the impact are negative. Furthermore, even when they are positive, compared to other proven interventions, their cost effectiveness is low. In Chile, Bellei (2009) estimates the impact of the Chilean full school day program, a policy that moved away public schools from half-day shifts to full school day shifts. The author uses difference-in-differences to estimate effect on test outcomes, and finds that the program increased the language achievement in 0.05-0.07 standard deviations and mathematics achievement in 0.12 standard deviations. In Uruguay, Cerdan-Infantes and Vermeersch (2007) evaluate the Uruguay full-time school program, a intervention that lengthened the school day from a half-day to full-day, and provided additional inputs to schools to make this possible. The results indicate that the program increased student test scores 0.063 standard deviations per year in mathematics and 0.044 standard deviations per year in languages. In Peru, Agüero (2016) estimates the impact of *La Jornada Escolar Completa* (The full-time schooling), a program that extends from 35 to 45 hours per week, together with pedagogic, management and support components. Using a RDD, the author finds that the program increases mathematics scores in 0.24 standard deviations and languages scores in 0.15 standard deviations. In Colombia, Bonilla (2011) uses a LATE to estimate the effect of full-time schooling in test scores. As an instrumental variable, the supply of the full-time schooling in the city is used, and an impact of 0.06 standard deviations on test scores of *Las Pruebas Saber 11* is found. Hincapie (2016), in its turn, uses a fixed effect approach to estimate the impact of

full-time schooling in test scores. The author finds that schools that adopted full-time have, in average, 0.05 to 0.16 standard deviations higher scores than half-day schools. Overall, effects of extension of school time and full-time schooling are significant, but mostly with a very small magnitude (ranging from 0.05 to 0.24 standard deviations).

In Brazil, the impact evaluations in extension of school day are mainly for the *Mais Educação*[9] Federal Program. Almeida et al. (2016) analyze a panel data of municipal characteristics, administrative data from the program, school census and students assessments from year 2007 to 2011 to estimate the impact of *Mais Educação* on tests scores and dropout rates. They use propensity score matching to deal with the potential problem of selection bias, together with a difference-in-differences approach, to compare control and treated groups before and after the implementation of the program. The authors don't find any positive impacts of the program on outcomes, such as Portuguese tests scores and drop out rates. Even more, they find negative impacts on mathematics test scores.

In a similar fashion, Oliveira and Terra (2016) use a RDD to estimate the effect of *Mais Educação*, and the results suggests that the program has no impact on educational outcomes. Pereira (2011) also estimates the impact of *Mais Educação* on drop out rates, approval rates, and test scores. Using only schools from *Minas Gerais*, from 2009 and 2010, the author applies the difference-in-difference methodology and finds impacts of the program on drop out rates, but does not find impact on approval rates and test scores. Gandra (2017), with a propensity score matching difference-in-difference estimator, finds no significant impact of *Mais Educação* on any educational outcome as well. Looking at a different program in Brazil, Aquino (2011) estimates the impact of *Escola de Tempo Integral* program, an intervention that extends school day in public schools in the state of São Paulo. Using propensity score matching and difference-in-differences estimators, the author finds that the program increases Portuguese test scores (0.13 standard deviations), but no significant effects were found for mathematics test scores nor for approval rates.

The literature of full-time schooling and school extension shows no consensus on the impact of more time in school. Some studies find small positive significant effects, others do not find any effect, and some find negative significant effects. One hypothesis is the fact that all the evaluated programs are not only extension of school time, but extension

[9] *Mais Educação* is, by the government definition, a "Ministry education's strategy to induce the construction of an agenda of full time schooling in state and municipal education networks that extends the school day in public schools for at least 7 hours a day through optional activities in macrofields: follow-up pedagogical; environmental education; sports and leisure; human rights in education; culture and arts; digital culture; health promotion; communication and use of media; research in the field of natural sciences and economic education" (http://portal.mec.gov.br/component/content/article?id=16689:saiba-mais-programa-mais-educacao). In short, it is a federal initiative to promote full-time schooling in public schools, with the objective of improving quality of education. To do so, extra financial support is provided directly to the schools that choose to participate in the program. The schools, in turn, have relatively autonomy to choose how to use this money among activities of the macro fields, and one subject from traditional fields, such as mathematics and Portuguese.

of school time used in a specific way. Extended day policies are usually mixed to other education policies, and heterogeneity in the implementation are relevant to the results, as it can be observed across different studies. If we assume that time is an important input to learning, but needs to interact with other variables, such as how time is spent, infrastructure or quality of teachers, the different ways that this extra time is used could have different impacts, even results with opposing signals. Our work tries to cast some light in this ambiguity. With the "full-time schooling program", we seek to: (i) estimate the impact of the program on educational outcomes, such as test scores and flow indicators; (ii) estimate other possible effects that the program has on other variables, such as scores on untreated schools and on composition of students in the schools; and (iii) understand the possible channels through the program may have impact on the outcomes. Moreover, this work's contribution is the evaluation of a program that has not been yet evaluated.

To achieve these goals, we use a difference-in-differences methodology, since treatment assignment is not random. Together with data on ENEM, *Prova Brasil*, school census, we estimate the impact of the program on educational outcomes. Additionally, to add robustness to our results, we perform matching approach, to address possible problems with the assumption of parallel trends.

Going further, we focus on possible impacts that the program could have in other variables, to understand some indirect effects of the program. We are going to estimate the impact of program on schools characteristics and on average student characteristics, using the same difference-in-differences methodology. We want to asses possible changes in school composition, and to check if the effects of the program are not resulting from the program itself, but from the change in students or teachers — a selection effect. Additionally, we are going to estimate the impacts of the program on the cities' educational outcomes average, excluding from the calculation of the mean schools that are treated. Here we seek to understand if treated schools are somehow attracting good students from other schools in the city — a "spillover effect".

Later, we use some variability in the way which the program was implemented in the schools and the differences between implementation in the beginning and in the end to isolate and identify the channels, using the same methodology as before.

This work is going to be divided as follows: Chapter 2 presents the "Full-time schooling program", introducing its history, characteristics and institutional aspects. In chapter 3 we analyze the data set and the methodology used in the work. In chapter 4, we can find the results of the estimations. Chapter 5 concludes this dissertation.

2 The Full-Time Schooling Program

The full-time schooling program is a public-private initiative that focus on improving the quality of public school in Brazil. To do this, the program implements a set of innovative pedagogic and management practices in the chosen schools[1]. It was first implemented in 2004 and since then it has been expanding to several states. The idealizer and creator of the program is a non-profit private entity that we are going to address as the "Institute". The Institute's mission is "to improve the quality of public education, through the mobilization of society, following the ethics of co-responsibility, producing innovative and replicable educational solutions in terms of content, method and management".

The pedagogic practices are centralized in the concept of the "Life project". In this concept, the school must provide for the student to develop fundamental conditions for his development, not only professional and cognitive, but also personal and social. Meaning must be provided for the student for him to develop his life project. For this to happen, the student must be seen as the centrality of the school project. The school, in its turn, must be directed by three main principles. First, Youth Protagonism, in which the young is responsible for his story, his condition, and for his own choices. Second, it the Pedagogy of Presence, in which the student must spend time with people who care about him, want to participate and influence in his grow. This includes not only teachers, but also other school employees. All adults in the school must serve as an example. Third, the education must not be lined only with the cognitive, but also with socio-emocional competences.

To sustain this school project, teacher's practices must be oriented by some pillars: (i) the formation of students values, by means of social experiences in the school, with other students and teachers; (ii) academic excellence, since without it, there is no meaning for the life project; (iii) formation of competences that allow the student to live fully in the 21th century.

In practical terms, the curriculum must evolve in a way to integrate the life project. More classes are needed in core disciplines, such as mathematics and Portuguese language, and in elective disciplines, related mainly with the life project classes. In effect, the full-time schooling is a natural consequence of the school project, that is, the necessity of extra time for this more demanding curriculum. In the end, school time is extended to 7.5h or 9h per day.

All these pedagogical practices are of no use if the results were not reasonable,

[1] The information about the program was based on interviews, on Magalhães (2008) and on Junior (2011)

perennial, and sustainable. Hence, a management model was created to sustain the school project and to ensure that all this pedagogic work was going to be translated into results. This model was based in management practices common in the business sector. These practices consist in:

- adoption of specific criteria for the selection of managers of schools;
- adoption of selection criteria and removal of teachers from the network;
- evaluation of managers, teachers (and ultimately of students);
- adoption of incentive systems for teachers, based on their own results and pupils' result; and
- adoption of a management instrument that focus in goal setting, plans and programs of action, creation of indicators, evaluation of results and problems' solving.

All these implementations needed to be supported by law in the state legislation. Hence, changes in existing laws and introduction of new laws and other legal instruments had to be done.

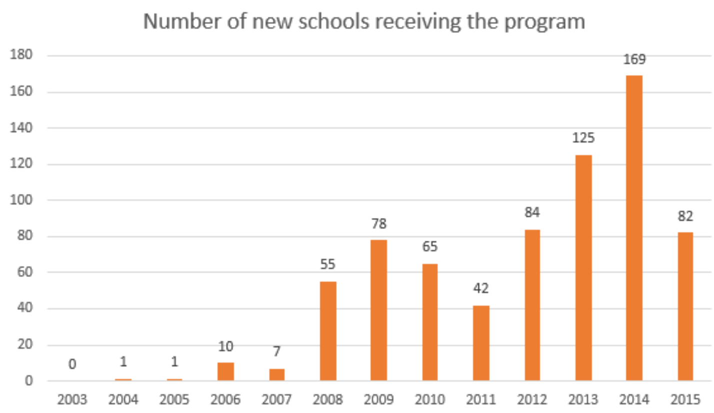
Initially, the proposal was to make a selection process for the students based on merit, measured by a test of Portuguese and mathematics. However, a intervention by the Public Prosecutor's Office (lit. *Ministério Público*) changed that the selection criteria to proximity between the school and the student's residence (Junior (2011)). In case of excess of demand, there was a additional criteria based on academic record. However, as Dias and Guedes (2010) pointed, that was only a formality, since the student's recorded score were not consistent with the student performance in the placement course. Also, since the inflation of grades was a common practice, student records usually didn't bear any relation with the student's knowledge (Magalhães (2008)).

The implementation of all these changes are conducted by the Secretary of education, which usually creates a executing body to implement the program in chosen schools. The decision of in which school the program is going to be implemented is conducted through studies made by the secretary of education and by the Institute. Among the factors that influence in the decision, we can highlight two: cities or a group of cities in a region with population of more or less 200 000; and support of the city prefecture in infrastructure and school bus of students.

The program itself is not fixed, and it is in constant process of evolution and improvement. Since its beginning, it has gone through some changes, but the core of the pedagogic and management stayed the same. The implementation of the program can be divided in two stages: the first stage of the program took place in the years from 2004 -

2007, in only one state. We are going to address it as early years stage. The second stage took place from 2008 to nowadays, and are going to address it as late years stage.

Figure 7 – New schools



Sources: Administrative data from the program.

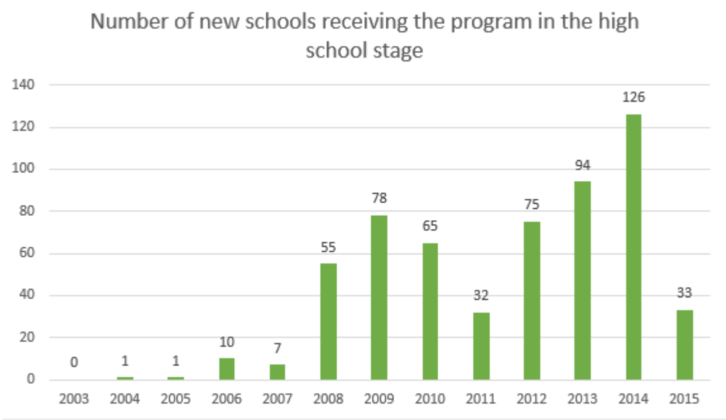
The main difference between these two stages is the Institute's role in the implementation. In the beginning of the program, the Institute did not enter in the management of schools, but it had a close relationship with the executing body. This allowed the creation and the testing of the practices idealized, with the backup of human resources and legal sectors in the secretary. Together with the managing body in the secretary of education, they helped to train employees, principals, teachers and implement all new methodologies in schools. The implementation of all envisioned practices were more "rigorous". Process of selection, incentive systems, evaluations were all implemented as they were initially thought. Also, they had an important role in making the necessary changes in the legal framework for the program to function adequately. In practical terms, the Institute always had a chair in the managing council of the participating schools, which allowed them to accompany the progress and make necessary changes in the participating schools.

Additionally, the Institute, in the early stage, contributed with funding for construction of libraries, laboratories, school lunch, uniforms, books and performance bonus for the teachers.

In the second stage of the program, the institute stepped out of any managing role, passing everything to the secretary of education. This way, the decision of implementing certain practices were up to the secretary, and, in the last instance, up to the governor. Natural polemic changes, such as the removal and evaluation of teachers, were often weaken or left. The Institute role could be summarized as a consultant role: when partnership starts, the Institute goes to the state government, trains secretary of education's employees, and transfer the "technology" to implement the program in schools. The governor and the secretary of state could ultimately choose not to follow some recommendations.

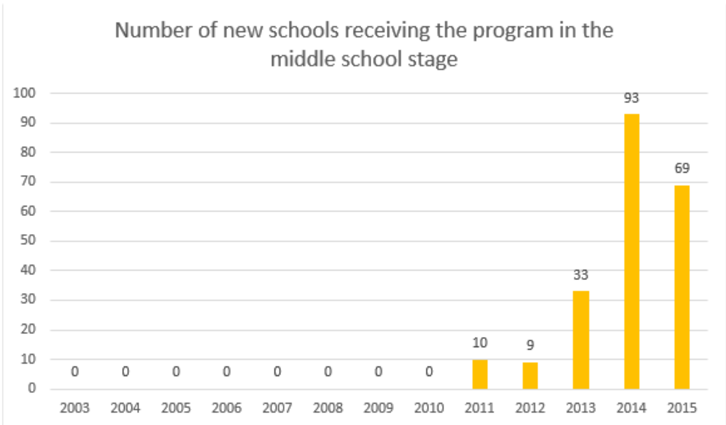
This way, the late years stage of the program were less "rigorous" in the implemen-

Figure 8 – New schools receiving the program in the high school



Sources:Administrative data from the program.

Figure 9 – New schools receiving the program in the middle school

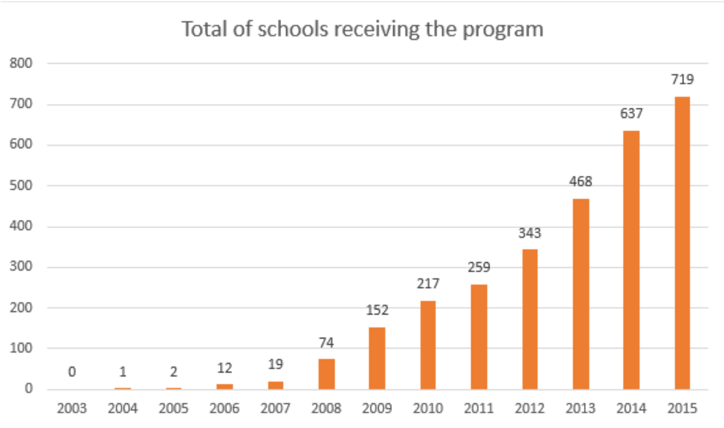


Sources:Administrative data from the program.

tation of the program in the schools. Nevertheless, core concepts of the program, such as "life project" and full-time schooling, were always implemented in the participating schools.

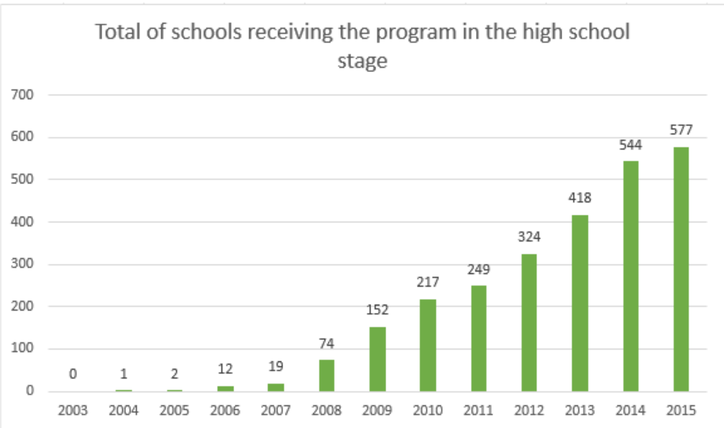
In figures 7, 8 and 9 we have plots of new schools entering in the program through years. For high schools, the program started in 2004, and its participating schools grew up to 577 in 2015 (figure 11). For middle schools, the program started in 2011, and the number of participating middle schools grew up to 214 (figure 12). As we can see, the number and variance through time of implementations of the program in middle schools is small. We are going to focus most of the estimations on high school schools. Nevertheless, we present the impact of the program for the middle school sample.

Figure 10 – Total of schools receiving the program



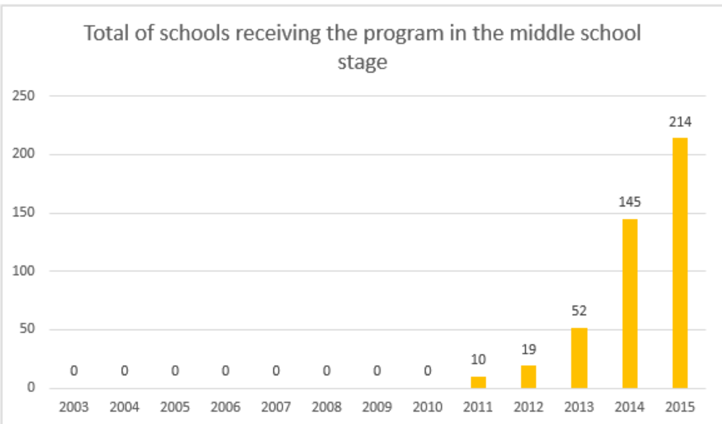
Sources: Administrative data from the program.

Figure 11 – Total of schools receiving the program in the high school



Sources: Administrative data from the program.

Figure 12 – Total of schools receiving the program in the middle school



Sources: Administrative data from the program.

3 Data and Empirical Strategy

3.1 Data

3.1.1 Main Samples

To estimate the impacts of the full-time schooling program, we are going to use data sets from different sources: ENEM, *Prova Brasil*, and the School Census.

ENEM, short for "National High School Exam" in Portuguese, is an optional annual exam that students can take in the end of high school. It was created in 1998 with the primary object of assessing the quality of high school education in Brazil. In 2009, ENEM changed its focus, and it started to be used as a criteria for admission in most of public universities and for scholarship grants in private universities. Since then, its importance has increased significantly. Enrollment in the exam is high, reaching 5 millions registered candidates in 2018, making one of the largest exams for college admission in the world.

The exam consists in multiple-choice and writing tests. Their results will be used as outcomes of interest in our estimations: for every student that took the ENEM exam, we have scores for their tests. Additionally, when doing the exam, candidates have to fill a socioeconomic questionnaire. Hence, information coming from this questionnaire is also going to be used. Specifically, data about mother and father's years of schooling, number of people in the household, family income and about participation of the student in the workforce.

Table 1 – Main variables

Database	Source	Variable	Use	Description
ENEM	INEP	ENEM writing test score	High school sample	School's average grades in ENEM writing test score, standardized in each year
		ENEM multiple choice test score	High school sample	School's average grades in ENEM Multiple Choice test score, standardized in each year
		Concluded High School in the same year as doing ENEM	High school sample	Percentage of students in the school who gradutaed from High School in the same year they did ENEM
		Lives with more than six people	High school sample	Percentage of students in the school who live with more than six people
		Father has college degree	High school sample	Percentage of students in the school whose father has a college degree
		Mother has college degree	High school sample	Percentage of sutdents in the school whose mother has a college degree
		Family Income is 5 minimum wages or higher	High school sample	Percentage of students in the school whose family income is 5 minimum wages or higher
		Looked or is looking for a job	High school sample	Percentage of students in the school who looked or are looking for a job
Prova Brasil	INEP	Prova Brasil Language test score	Middle school sample	School's average grades in Prova Brasil language test score - 9th grade
		Prova Brasil Mathematics test tcore	Middle school sample	School's average grades in Prova Brasil Mathematics test score - 9th grade
		Mother has college degree	Middle school sample	Percentage of students in the school whose mother has a college degree
		Father has college degree	Middle school sample	Percentage of students in the school whose father has a college degree
School Census	INEP	Promotion Rate	Both samples	Percentage of students who passed the grade, standardized in each year
		Failure Rate	Both samples	Percentage of students who failed to pass the grade, standardized in each year
		Drop-out Rate	Both samples	Percentage of students who drop out of school, standardized in each year
		Grade-age Distortion Rate	Both samples	Percentage of students, in each grade, whose age is higher that the recommendad age
		Rural Area	Both samples	Dummy variable, equals one if schools are in Rural Area
		Building	Both samples	Dummy variable, equals one if school's place of operation is a building
		Principal's office	Both samples	Dummy variable, equals one if school has a principal's office
		Teacher's room	Both samples	Dummy variable, equals one if school has a teacher's room
		Library	Both samples	Dummy variable, equals one if school has a library
		Internet Access	Both samples	Dummy variable, equals one if school has internet
		Waste Collection	Both samples	Dummy variable, equals one if school has waste collection
		Electricity	Both samples	Dummy variable, equals one if school has access to the public electricity network
		Water	Both samples	Dummy variable, equals one if school has access to the public water supply network
		Sewage	Both samples	Dummy variable, equals one if school has access to the public sewage system
		Computer Laboratory	Both samples	Dummy variable, equals one if school has a computer laboratory
		Science Laboratory	Both samples	Dummy variable, equals one if school has a science laboratory
		Sports Court	Both samples	Dummy variable, equals one if school has a sports court
		Number of Students in High School	High school sample	Total number of students in high school
		Number of Female Students in High School	High school sample	Total number of female students in high school
		Number of Students in Middle School	Middle school sample	Total number of students in middle school
		Number of Female Students in Middle School	Middle school sample	Total number of female students in middle school
		Number of White Students in Middle School	Middle school sample	Total number of white students in middle school
		Number of Classes in Middle School	Middle school sample	Total number of classes in middle school
		Number of Teachers in High School	High school sample	Total number of teachers in high school
		Number of Teachers with High School Degree in High School	High school sample	Total number of teachers with high school degree in high school
		Number of Teachers with College Degree in High School	High school sample	Total number of teachers with college degree in high school
		Proportion of High School Teachers with College Degree	High school sample	Proportion of high school teachers with college degree
		Number of Teachers in Middle School	Middle school sample	Total number of teachers in middle school
		Number of Teachers with College Degree in High School	High school sample	Total number of teachers with college degree in middle school
		Propotion of Middle School Teachers with College Degree	Middle school sample	Proportion of middle school teachers with college degree
IBGE		GDP per Capita	Both samples	Gross Domestic Production divided by population
		City Population	Both samples	City population
MEC/FNDE		Mais Educação Federal Program	Both samples	Dummy variable, equals one if schools adhered to Mais Educação
Administrative Data From The Full-time schooling Program		Treatment dummy	Both samples	Dummy variable, equals one if, in a given year, school is being treated
		Middle School dummy	Both samples	Dummy variable, equas one if school receive program in middle school
		High School dummy	High school sample	Dummy variable, equals one if school receive program in high school
		Mediator 1	High school sample	Dummy variable, equals one if state governor was engaged with the implementation of the program
		Mediator 2	High school sample	Dummy variable, equals one if secretary of education was engaged with the implementation of the program
		Mediator 3	High school sample	Dummy variable, equals one if there was a dedicated team in secretary of education for the implementation of the program
		Mediator 4	High school sample	Dummy variable, equals one if changes in the legal framework took place
		Mediator 5	High school sample	Dummy variable, equals one if all necessary changes in the legal framework took place, by the prescribed deadline
		Mediator 6	High school sample	Dummy variable, equals one if a process of selection and removal of principals took place
		Mediator 7	High school sample	Dummy variable, equals one if a process of selection and removal of teachers took place
		Mediator 8	High school sample	Dummy variable, equals one if the life project was implemented in the curriculum

Prova Brasil, in its turn, is the exam that assess the quality of middle school education. Created in 2005, it evaluates the development of 5^{th} and 9^{th} grade students. This multiple-choice exam is done in all public middle schools with more than 20 students every other year. Like in ENEM, students have to fill a questionnaire about socioeconomic, cultural, habits and studying environment. This way, we have information about scores, about the school and some characteristics of the student that did the exam. In our case, we use the test scores for the 9^{th} grade of middle school.

The above mentioned two exams are conducted by INEP ("*Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*"), an agency controlled by the Ministry of Education. We aggregate the data into school level, so we can control for school characteristics, if needed. For the educational outcomes of flow indicators, INEP provides them in school level.

We also use the School Census, an annual survey of all students enrolled in any school in Brazil. It is an extensive survey of all students, from primary to high school, containing a different range of questions about the teachers, the school, and the students itself. In particular, we exploit school level data: school characteristics and number of enrollments adn teachers, which are going to be used as controls in the estimation.

In addition, we control for other characteristics, such as city characteristics — GDP per capita and population, which can be found in IPEADATA [1]—, and participation in the federal program *Mais Educação*, whose information was provided by the administritative data from MEC/FNDE[2].

Finally, we have at our disposal administrative data from the full-time schooling program. In this data-set, we have information about which schools joined the program, when they joined — different schools entered the program in different years—,and some heterogeneity about how the program was implemented in the school.

In the end, we have two main samples: one composed of high schools and another composed of middle schools. In both samples, we focus only in public schools: Schools of the treatment group are schools that participate in the program, and schools of the control group are public schools of participating states. In the high school sample, we have observations in school level, containing the outcome variables (ENEM test scores and flow indicators), school characteristics (School census), average characteristics of students in the school (ENEM), city characteristics (GDP per capita and total population), and information from the program, which includes treatment indicator and some variance between implementation. Since the program started in 2004, we have observations ranging

[1] *Instituto de Pesquisa Econômica*(IPEA) is a federal public agency whose research activities provide technical and institutional for government policy. IPEADATA is its data system, which contains data about many indicators and variables in Brazil.

[2] Ministry of Education and Culture (MEC) and National Fund of development and Education (FNDE)

from the year 2003 to 2015 [3]. Therefore, we have a panel data structure. Note that some schools don't have observations in all the years: the resulting sample is an unbalanced panel. For the middle school sample, we have observations in school level, containing the the outcome variables (*Prova Brasil* Test Scores and flow indicators), schools characteristics (School census), city characteristics (GDP per capita and total population), and information from the program. For middle schools, the program started in 2011, so the sample contains observations from 2007 to 2015. As in the high school sample, we have an unbalanced panel data structure. In table 1 we have a list of the main variables used in the estimations, their descriptions, their source and the sample where they are used.

3.1.2 Descriptive Statistics

In table 2 and table 3 we have the descriptive statistics on the main variables for high school and middle school samples. The high school sample is composed of 152924 observations, ranging from 2003 to 2015. The middle school sample is composed of 56471 observations, ranging from 2007 to 2015. Both are unbalanced panels, since the number of schools increases from year to year. New schools are entering the samples or not leaving them. For example, there are some schools whose students did not do the ENEM test in a given year, and there are no information coming from ENEM about the students. Ideally, we would want a balanced panel, keeping only schools whose information are present in all 11 years. However, this unbalance is seen also in participating schools: some of them appear only in late years of the sample.

Tables 4 and 5 report the descriptive statistics of both samples in three different years. For each year reported, we have the number of schools in the sample, the variable mean and its standard deviation. In the high school sample, we have 9182 schools in 2004, 12980 in 2009 and 13217 schools in 2015. In the middle school sample, we have 11649 schools in 2011, 11588 schools in 2013 and 10996 schools in 2015.

In the high school sample, when analyzing average students characteristics obtained through ENEM, we have that the percentage of students who lives with more than six people decreased from 7% to 2%, more or less. Likewise, the percentage of students whose family income is 5 minimum wages or higher has decreased, from 0.316 to 0.172. Approximately 0% of the schools belong to the treatment in 2004; in 2015 4% of the schools are being treated. Looking at school characteristics obtained through the School Census, we have that the average number of students in high school has decreased from 2004 to 2015, going from 434.6 to 392.7. Similarly, in the middle school sample, we have that the percentage with students whose father or mother has college degree increased from 2011 to 2015 (0.0716 to 0.0999 for mother and 0.0643 to 0.0886 for father). Approximately 0.085%

[3] The available data from the program ranges from 2004 to 2015, but the program is still active nowadays.

of the schools belong to the treatment in 2004; in 2015 1.72% of the schools are being treated. The number of students have decreased from 194.9 to 164.2.

Tables 6 and 7 reports the balance between schools that will participate in the Full-time schooling program and non-participating schools, in the first available year of each sample, for school and student characteristics, as well as outcomes of interest. As we can see, we can reject the equality of means for most of the variables and outcomes between treatment and control groups. Indeed, this was expected, since the program is not randomly assigned. In the high school sample, schools selected to the program have lower average scores in the ENEM multiple choice test and in the ENEM writing test. Conversely, in the middle school sample, schools selected to the program have higher average scores in the *Prova Brasil* language and mathematics tests. To a certain extent, this was unexpected. Since the program is not random, and the choice of new schools are left to the secretary of education, one would expect that low performance schools were to be chosen.

Table 2 – Descriptive statistics - High school sample - All years

Variable	Nobs	Mean	sd
ENEM multiple choice test score	149372	0.161	1.022
ENEM writing test score	151691	0.0688	0.975
Promotion Rate	96745	-0.0177	0.984
Failure Rate	96745	0.0121	0.975
Drop-out Rate	96745	0.0152	0.984
Grade-age Distortion Rate	95730	-0.00281	0.983
Concluded High School in the same year as doing ENEM	152024	57.88	64.38
Lives with more than six people	149824	0.0584	0.0903
Father has college degree	149100	0.145	0.213
Mother has college degree	149538	0.171	0.224
Family Income is 5 minimum wages or higher	149804	0.199	0.27
Looked or is looking for a job	148878	0.531	0.314
Rural Area	152024	0.0352	0.184
Building	151772	0.989	0.105
Principal's office	151934	0.951	0.216
Teacher's room	151933	0.957	0.204
Library	151933	0.621	0.485
Computer Laboratory	151931	0.786	0.41
Science Laboratory	151931	0.441	0.496
Sports Court	151931	0.772	0.42
Internet Access	150654	0.934	0.248
Waste Collection	151934	0.986	0.115
Electricity	151934	0.999	0.0281
Water	151934	0.959	0.199
Sewage	151933	0.828	0.378
Number of Teachers in High School	152024	77.39	89.61
Number of Teachers with High School Degree in High School	152024	2.273	7.173
Number of Teachers with College Degree in High School	152024	75.09	87.83
Proportion of High School Teachers with College Degree	134555	0.961	0.104
Mais Educação Federal Program	152024	0.34	0.474
Number of Students in High School	134040	351.1	335
Number of Female Students in High School	133810	184.7	179.5
City Population	152024	2032309	3688121
GDP per Capita	152024	33095	25684

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Descriptive Statistics of the main variables in the high school sample, in all years. The outcome variables — ENEM multiple choice test score, ENEM writing test score, Promotion rate, Failure rate, Drop-out rate, and Grade-age distortion rate — are standardized in each year. However their means and standard errors are not 1 and 0, respectively. That happens because the standardization of these variables was made before dropping out observations which have missings in other data bases.

Table 3 – Descriptive statistics - Middle school sample - All years

Variable	Nobs	Mean	sd
Prova Brasil Mathematics Test Score	56471	0.0792	0.905
Prova Brasil Language Test Score	56471	0.0427	0.895
Promotion Rate	50344	-0.236	0.961
Failure Rate	56428	0.213	0.955
Drop-out Rate	56428	0.164	0.924
Grade-age Distortion Rate	56416	0.194	0.903
Mother has college degree	56336	0.0722	0.0656
Father has college degree	56308	0.0638	0.0623
Rural Area	56471	0.0919	0.289
Building	56446	0.998	0.0474
Principal's office	56471	0.942	0.234
Teacher's room	56471	0.929	0.256
Library	56471	0.529	0.499
Computer Laboratory	56471	0.854	0.353
Science Laboratory	56471	0.251	0.433
Sports Court	56471	0.765	0.424
Internet Access	55801	0.881	0.324
Waste Collection	56471	0.966	0.182
Electricity	56471	0.999	0.0263
Water	56471	0.929	0.257
Sewage	56471	0.718	0.450
Number of Class Rooms in use	56471	12.63	6.307
Number of Teachers in Middle School	56471	102.1	60.88
Number of Teachers in High School	56471	60.63	84.73
Number of Teachers with College Degree in High School	56471	96.82	60.37
Propotion of Middle School Teachers with College Degree	56041	0.931	0.122
Mais Educação Federal Program	56471	0.700	0.458
Number of Classes in Middle School	56471	12.20	6.406
Number of Students in Middle School	56471	392.2	240.1
Number of Female Students in Middle School	56471	193.3	119.7
Number of White Students in Middle School	56471	116.7	121.6
City Population	56471	1626000	3470000
GDP per Capita	56471	31374	27651

Source: School Census, *Prova Brasil*, INEP, IBGE, MEC/FNDE, Administrative Data. *Note:* Descriptive Statistics of the main variables in the high school sample, in all years. The outcome variables — *Prova Brasil* Mathematics Test Score, *Prova Brasil* language test score, Promotion rate, Failure rate, Drop-out rate, and Grade-age distortion rate — are standardized in each year. However their means and standard errors are not 1 and 0, respectively. That happens because the standardization of these variables was made before dropping out observations which have missings in other data bases.

Table 4 – Descriptive statistics - High school sample - 2004 - 2009 - 2015

	2004			2009			2015		
Variable	Nobs	Mean	sd	Nobs	Mean	sd	Nobs	Mean	sd
ENEM multiple choice test score	9182	0.163	1.025	12303	0.172	1.035	12950	0.198	1.001
ENEM writing test score	9182	0.118	0.970	12980	-0.0659	0.975	13216	0.0790	0.968
Promotion Rate				10427	-0.0181	0.990	11550	-0.00527	0.974
Failure Rate				10427	0.0174	0.990	11550	0.00212	0.965
Drop-out Rate				10427	0.00967	0.988	11550	0.00747	0.965
Grade-age Distortion Rate				10312	0.00147	0.991	11558	-0.00394	0.974
Concluded High School in the same year as doing ENEM	9182	64.22	65.94	12980	45.73	54.80	13217	58.86	65.91
Lives with more than six people	9015	0.0786	0.0962	12146	0.109	0.140	13217	0.0202	0.0518
Father has college degree	9005	0.147	0.215	12052	0.145	0.221	13216	0.144	0.203
Mother has college degree	9013	0.157	0.215	12127	0.179	0.237	13216	0.186	0.223
Family Income is 5 minimum wages or higher	9017	0.316	0.308	12143	0.212	0.279	13206	0.172	0.244
Looked or is looking for a job	9013	0.659	0.277	12131	0.653	0.294	13216	0.367	0.283
Rural Area	9182	0.0270	0.162	12980	0.0352	0.184	13217	0.0423	0.201
Building	9163	0.987	0.111	12980	0.984	0.124	13055	1	0
Principal's office	9163	0.960	0.195	12980	0.943	0.233	13217	0.945	0.227
Teacher's room	9162	0.963	0.190	12980	0.942	0.234	13217	0.952	0.214
Library	9162	0.658	0.474	12980	0.543	0.498	13217	0.549	0.498
Computer Laboratory	9160	0.0830	0.276	12980	0.832	0.374	13217	0.879	0.326
Science Laboratory	9160	0.275	0.447	12980	0.412	0.492	13217	0.466	0.499
Sports Court	9160	0.375	0.484	12980	0.772	0.419	13217	0.810	0.392
Internet Access	8950	0.857	0.350	12980	0.944	0.229	13217	0.961	0.193
Waste Collection	9163	0.981	0.136	12980	0.985	0.121	13217	0.988	0.108
Electricity	9163	0.999	0.0330	12980	0.998	0.0392	13217	1.000	0.0213
Water	9163	0.959	0.198	12980	0.957	0.203	13217	0.957	0.203
Sewage	9162	0.827	0.378	12980	0.820	0.385	13217	0.831	0.375
Number of Teachers in High School	9182	23.58	16.13	12980	88.80	96.19	13217	101.4	98.40
Number of Teachers with High School Degree in High School	9182	1.058	3.740	12980	3.311	9.407	13217	5.483	9.963
Number of Teachers with College Degree in High School	9182	22.51	16.15	12980	85.44	93.38	13217	95.87	94.85
Proportion of High School Teachers with College Degree	8946	0.947	0.155	10608	0.954	0.0949	11606	0.934	0.101
Treatment dummy	9182	0	0	12980	0.00955	0.0973	13217	0.0415	0.200
Mais Educação Federal Program	9182	0.327	0.469	12980	0.331	0.470	13217	0.428	0.495
Number of Students in High School	8939	434.6	402.0	10558	332.8	324.7	11548	302.7	276.8
Number of Female Students in High School	8915	230.4	218.8	10551	176.7	174.3	11539	155.4	143.5
City Population	9182	2012000	362000	12980	2073000	3673000	13217	2115000	3863000
GDP per Capita	9182	25125	20203	12980	32972	23603	13217	35341	24248

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Descriptive Statistics of the main variables in the high school sample, in all years. The outcome variables — ENEM multiple choice test score, ENEM writing test score, Promotion rate, Failure rate, Drop-out rate, and Grade-age distortion rate — are standardized in each year. However their means and standard errors are not 1 and 0, respectively. That happens because the standardization of these variables was made before dropping out observations which have missings in other data bases. Flow variables — Promotion rate, Failure rate, Drop-out rate, and Grade-age distortion rate — were not available until 2007. Hence, in 2004, there are no information about them.

Table 5 – Descriptive statistics - Middle school sample - 2011 - 2013 - 2015

	2011			2013			2015		
Variable	Nobs	Mean	sd	Nobs	Mean	sd	Nobs	Mean	sd
Prova Brasil Mathematics Test Score	11649	0.0696	0.898	11588	0.139	0.892	10996	0.201	0.864
Prova Brasil Language Test Score	11649	0.00423	0.872	11588	0.133	0.889	10996	0.209	0.895
Promotion Rate	11645	-0.226	0.946	11588	-0.219	0.971	4911	-0.179	0.967
Failure Rate	11645	0.209	0.969	11588	0.173	0.984	10995	0.203	0.950
Drop-out Rate	11645	0.130	0.852	11588	0.192	0.890	10995	0.176	0.866
Grade-age Distortion	11639	0.197	0.883	11583	0.176	0.896	10996	0.216	0.905
Mother has college	11649	0.0716	0.0644	11537	0.0821	0.0678	10938	0.0999	0.0756
Father has college degree	11649	0.0643	0.0621	11537	0.0735	0.0657	10938	0.0886	0.0735
Rural Area	11649	0.116	0.320	11588	0.108	0.311	10996	0.104	0.305
Building	11649	0.998	0.0404	11588	0.996	0.0594	10971	1	0
Principal's office	11649	0.958	0.201	11588	0.937	0.244	10996	0.933	0.251
Teacher's room	11649	0.922	0.268	11588	0.929	0.256	10996	0.942	0.234
Library	11649	0.458	0.498	11588	0.448	0.497	10996	0.447	0.497
Computer Laboratory	11649	0.902	0.298	11588	0.923	0.267	10996	0.918	0.275
Science Laboratory	11649	0.260	0.439	11588	0.266	0.442	10996	0.269	0.443
Sports Court	11649	0.736	0.441	11588	0.776	0.417	10996	0.796	0.403
Internet Access	11523	0.911	0.284	11535	0.916	0.278	10996	0.940	0.237
Waste Collection	11649	0.958	0.201	11588	0.967	0.180	10996	0.972	0.166
Electricity	11649	0.999	0.0334	11588	1.000	0.0208	10996	1.000	0.0213
Water	11649	0.919	0.272	11588	0.921	0.270	10996	0.927	0.261
Sewage	11649	0.699	0.459	11588	0.707	0.455	10996	0.728	0.445
Number of Class Rooms in use	11649	12.36	6.614	11588	12.86	6.321	10996	12.84	6.438
Number of Teachers in Middle School	11649	106.4	63.27	11588	102.3	60.88	10996	92.50	53.14
Number of Teachers in High School	11649	61.52	86.70	11588	61.44	87.05	10996	60.96	86.98
Number of Teachers with College Degree in High School	11649	102.0	62.93	11588	99.16	60.26	10996	86.21	51.94
Propotion of Middle School Teachers with College Degree	11573	0.944	0.108	11532	0.962	0.0832	10846	0.916	0.130
Treatment dummy	11649	0.000858	0.0293	11588	0.00431	0.0655	10996	0.0172	0.130
Mais Educação Federal Program	11649	0.709	0.454	11588	0.712	0.453	10996	0.703	0.457
Number of Classes in Middle School	11649	12.26	6.429	11588	11.97	6.242	10996	11.28	5.767
Number of Students in Middle School	11649	396.1	240.9	11588	372.5	225.3	10996	336.1	197.5
Number of Female Students in Middle	11649	194.9	119.8	11588	182.3	111.4	10996	164.2	97.74
Number of White Students in Middle	11649	121.5	125.8	11588	123.3	121.2	10996	118.0	104.4
City Population	11649	1591000	3434000	11588	1656000	3582000	10996	1715000	3666000
GDP per Capita	11649	33041	31094	11588	34383	30761	10996	32676	26443

Source: School Census, *Prova Brasil*, INEP, IBGE, MEC/FNDE, Administrative Data. *Note:* Descriptive Statistics of the main variables in the high school sample, for 2011, 2013, and 2015. The outcome variables — *Prova Brasil* Mathematics Test Score, *Prova Brasil* language test score, Promotion rate, Failure rate, Drop-out rate, and Grade-age distortion rate — are standardized in each year. However their means and standard errors are not 1 and 0, respectively. That happens because the standardization of these variables was made before dropping out observations which have missings in other data bases.

Table 6 – Sample balancing before matching - Average school characteristics in 2003, by treatment status - High school sample

year = 2003 Variable	Nonparticipating Schools		Schools that will participate in the Full-time Schooling Program		Difference (1)-(3)	
	Mean	sd	Mean	sd	Mean Diff	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
ENEM writing test score	0.05	1.01	-0.49	0.55	0.54***	(18.83)
ENEM multiple choice test score	0.15	1.05	-0.51	0.44	0.66***	(27.59)
Concluded High School in the same year as doing ENEM	72.01	75.78	96.40	81.91	-24.39***	(-5.98)
Lives with more than six people	0.08	0.09	0.14	0.10	-0.06***	(-12.94)
Father has college degree	0.14	0.21	0.03	0.04	0.11***	(35.63)
Mother has college degree	0.13	0.20	0.03	0.04	0.10***	(32.44)
Family Income is 5 minimum wages or higher	0.32	0.30	0.10	0.12	0.21***	(31.35)
Looked or is looking for a job	0.68	0.27	0.81	0.11	-0.12***	(-20.13)
Rural Area	0.02	0.15	0.01	0.08	0.02***	(3.55)
Building	0.99	0.11	0.99	0.10	-0.00	(-0.70)
Principal's office	0.96	0.20	0.93	0.26	0.03*	(2.53)
Teacher's room	0.96	0.19	0.95	0.22	0.01	(1.34)
Library	0.84	0.37	0.84	0.37	-0.00	(-0.08)
Internet Access	0.81	0.39	0.83	0.38	-0.01	(-0.57)
Waste Collection	0.98	0.13	0.96	0.19	0.02*	(2.30)
Electricity	1.00	0.01	1.00	0.00	-0.00	(-1.41)
Water	0.96	0.19	0.96	0.20	0.00	(0.28)
Sewage	0.82	0.39	0.75	0.43	0.06**	(2.90)
Computer Laboratory	0.67	0.47	0.56	0.50	0.12***	(4.75)
Science Laboratory	0.49	0.50	0.41	0.49	0.08**	(3.26)
Sports Court	0.82	0.39	0.74	0.44	0.07**	(3.27)
Number of Students in High School	434.57	404.75	635.79	370.29	-201.22***	(-10.73)
Number of Female Students in High School	226.44	213.68	342.18	202.64	-115.73***	(-11.30)
GDP per Capita	24783.67	19819.56	16213.72	17126.79	8569.96***	(9.94)
City Population	2067810.11	3632333.83	804528.50	2184003.71	1263281.61***	(11.14)
Mais Educação Federal Program	0.32	0.47	0.43	0.50	-0.11***	(-4.45)
Number of Teachers in High School	21.33	16.37	26.04	12.86	-4.70***	(-7.22)
Number of Teachers with High School Degree in High School	1.43	4.68	1.17	4.31	0.26	(1.19)
Number of Teachers with College Degree in High School	19.89	16.47	24.87	13.26	-4.98***	(-7.42)
Proportion of High School Teachers with College Degree	0.92	0.20	0.95	0.16	-0.03***	(-4.11)
Observations	8969		419		9388	

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Table reports average of main variables for the high school sample, in the first year of the data base. Columns (1) and (3) are the variable's means for non-participating schools and participating schools respectively, and (2) and (4) are their respective standard errors. Column (5) reports the difference between the means (columns (1) - (3)). Column (6) reports the p-value of the t-test to detect significant differences between the two groups. *significant at 10%. **significant at 5%. ***significant at 1%.

Table 7 – Sample balancing before matching - Average school characteristics in 2007, by treatment status - Middle school sample

year = 2007	Nonparticipating Schools		Schools Participating that will participate in the Full-time Schooling Program		Difference (1)-(3)	
Variable	Mean	sd	Mean	sd	Mean Diff	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
Prova Brasil Language Test Score	-0.05	0.95	0.39	0.78	-0.44***	(-7.20)
Prova Brasil Mathematics Test Score	-0.11	0.91	0.28	0.73	-0.39***	(-6.84)
Mother has college degree	0.05	0.05	0.07	0.05	-0.02***	(-3.82)
Father has college degree	0.05	0.05	0.08	0.05	-0.03***	(-7.02)
Rural Area	0.01	0.10	0.00	0.00	0.01***	(10.69)
Building	1.00	0.06	1.00	0.00	-0.00***	(-5.84)
Principal's office	0.94	0.24	0.96	0.20	-0.02	(-1.37)
Teacher's room	0.94	0.24	0.99	0.11	-0.05***	(-5.75)
Library	0.84	0.36	0.92	0.27	-0.08***	(-3.63)
Internet Access	0.80	0.40	0.98	0.13	-0.19***	(-16.49)
Waste Collection	0.98	0.13	1.00	0.00	-0.02***	(-13.15)
Electricity	1.00	0.01	1.00	0.00	-0.00	(-1.41)
Water	0.96	0.19	0.99	0.08	-0.03***	(-4.95)
Sewage	0.77	0.42	0.96	0.19	-0.19***	(-12.60)
Computer Laboratory	0.71	0.45	0.93	0.26	-0.22***	(-10.59)
Science Laboratory	0.22	0.41	0.40	0.49	-0.19***	(-4.82)
Sports Court	0.77	0.42	0.93	0.26	-0.16***	(-7.90)
Number of Class Rooms in use	12.72	6.04	13.57	4.61	-0.85*	(-2.35)
GDP per Capita	27987.57	24924.56	36174.21	18445.79	-8186.64***	(-5.64)
City Population	1615570.04	3328147.47	2400297.94	3756273.29	-784727.90**	(-2.67)
Mais Educação Federal Program	0.67	0.47	0.63	0.48	0.04	(1.03)
Number of Classes in Middle School	13.09	6.74	12.76	5.54	0.33	(0.76)
Number of Students in Middle School	447.85	264.74	437.19	225.71	10.66	(0.60)
Number of Female Students in Middle School	222.90	132.82	212.13	111.28	10.77	(1.23)
Number of White Students in Middle School	104.03	130.81	180.99	107.34	-76.97***	(-9.13)
Number of Teachers in Middle School	101.56	61.63	119.70	41.00	-18.15***	(-5.60)
Number of Teachers in High School	59.18	78.38	60.99	65.08	-1.81	(-0.35)
Number of Teachers with College Degree in High School	95.36	61.11	116.58	42.13	-21.22***	(-6.38)
Propotion of Middle School Teachers with College Degree	0.92	0.13	0.97	0.09	-0.05***	(-6.81)
Observations	10284		166		10450	

Source: School Census, *Prova Brasil*, INEP, IBGE, MEC/FNDE, Administrative Data. *Note:* Table reports average of main variables for the middle school sample, in the first year of the data base. Columns (1) and (3) are the variable's means for non-participating schools and participating schools respectively, and (2) and (4) are their respective standard errors. Column (5) reports the difference between the means (columns (1) - (3)). Column (6) reports the p-value of the t-test to detect significant differences between the two groups. *significant at 10%. **significant at 5%. ***significant at 1%.

3.2 Empirical Strategy

3.2.1 Baseline Empirical Model: Dynamic Differences-in-Differences

Our main objective is to estimate the impact of the Full-time Schooling Program on different outcomes. Given the structure of our samples — a panel data, with schools beginning treatment in different years —, our proposed methodology is the dynamic difference-in-differences, first seen in Autor (2003), where the author exploited the different timing of adoption of common law exception in states to estimate the causal impact of the exception on growth in Temporary Help Services and outsourcing. Let i be the i-th school, and t be a year between 2004 to 2015. Our estimation model is:

$$y_{it} = \beta X_{it} + \theta treatment_{it} + S_i + T_t + \varepsilon_{it} \tag{3.1}$$

where y_{it} is a outcome of interest (test scores for both ENEM and *Prova Brasil* and flow indicators (promotion, failure, dropout and grade-age distortion rates)), X_{it} is a vector of controls for school i in the year t, $treatment_{it}$ is a dummy that indicates that school i in year t is being treated, that is, if the program is implemented in this year or if it was implemented before, in this school. S_i are the school fixed effects, T_t are year fixed effects, and ε_{it} are unobserved terms. Ideally, the errors are to be clustered in states, since the implementation of the program were made mainly within each state secretary. Each state had a idiosyncrasy in its implementation. However, we have too few states in our sample, which can lead to problems in inference. Hence, we cluster errors in mesoregions [4]. We also cluster in states, in state years and in cities, in order to provide robustness for our results.

With the estimation model specified above, our coefficient of interest is θ. This parameter indicates the average of impacts of the program, through the years, on the dependent variables. The use of a difference-in-difference approach is adequate in this case because even though the treatment is not random in the school level, the approach removes biases resulting from permanent counfounders between treated and control groups and biases that are results of non-outcome trends. The key assumption for identification of θ is that trends in the outcomes would have been the same in schools that received the program and in schools that did not receive the program. That is, in the absence of treatment, the difference between the treatment and the control groups are constant over time. Since there are multiple periods and the change in treatment occurs in different years, it is not possible to provide a visual inspection of evolution of outcomes trends in years, for those schools which will receive treatment vis-à-vis those which will never receive treatment. However, we can assess the parallel trends hypothesis is some ways.

[4] Mesoregion is an area that is defined by its regional identity, constructed by the society living there.

First, we can allow for leads and lags of the treatment and test if the leads' coefficients estimates are non significant. That is, we can estimate the following equation:

$$y_{it} = \beta X_{it} + \sum_{m=0}^{M} \alpha_{-m} D_{i,-m} + \sum_{k=1}^{K} \alpha_{+k} D_{i,k} + S_i + T_t + \varepsilon_{it} \tag{3.2}$$

where $D_{i,-m}$ is the interaction between the dummy of year $-m$ relative to the entrance of treatment and the dummy that indicates that i was or will be treated, for a given school i. $D_{i,+k}$ is analogous for $+k$. For example, $D_{i,-1}$ assumes 1 if the school i is in the year before the treatment and it will be treated in the future (and 0 otherwise); $D_{i,0}$ assumes 1 if the school i is in the year of the treatment (and 0 otherwise); and $D_{i,+1}$ assumes 1 if the school i is in the year following the year of the treatment. We have, therefore, M leads and K lags of the treatment effect, in place of the single treatment effect, as in 3.1. Accordingly, the assessment of the parallel trends assumption is the test of $\alpha_j = 0$ for $\forall j < 0$. We estimate equation 3.2, and check if the coefficient estimates are insignificant. If this is the case, we have evidence that the differences in differences assumption should not be rejected.

Second, we can include time trends for treatment and control groups as regressors. To do so, we are going to include interactions between the dummy that indicates that a given school was or is going to be treated and the years dummies. Since most of schools will began treatment in the later years of the sample, these interactions are going to control for possible differences in pre-treatement trends of the control and treated groups. If the treatment correlate with these trends, the inclusion of the interactions will make the estimate of our parameter of interest not significant. On the contrary, if the estimate is significant, we have evidence that the treatment is not correlated to those trends, and that the estimation of the parameter of interest in 3.1 could be the average causal impact of the program in the educational outcomes.

3.2.2 Controlling for Observed Heterogeneity and for Pre-treatment Outcomes

As previously seen, we have that treatment is not random. The mechanisms of assignment of treatment status are somewhat arbitrary, since the institute and the secretary of education in the different states choose the cities and schools which will receive the program depending on political, educational, or economic reasons.

Potential threats to identification can arise from the non random assignment of the treatment. Despite the fact that a difference-in-differences approach removes bias from permanent confounders, the parallel trends hypothesis may be implausible if pre-treatment characteristics that are thought to be associated with the dynamics of the outcome variable are unbalanced between the treated and the untreated (Abadie (2005)),

that is, if the distribution of observable characteristics is different between the two groups. Lack of balance creates some problems: it can make inferences sensitive to ostensibly minor changes in the methods and specifications used and can make inferences imprecise (Imbens and Rubin (2015)). Therefore, we use a matching technique to select a sample where the treatment and control samples are more balanced than in the original sample.

Formally, when pursuing balance, we are in fact pursuing the plausibility of unconfoundedness hypothesis[5]:

$$(Y_i(0), Y_i(1)) \coprod treatment_i | X$$

where $Y_i(0)$ indicates the outcome Y for school i when not treated and $Y_i(1)$ indicates the outcome Y for school i when treated. The unconfoundedness hypothesis states that, conditional on observables, the treatment assignment is random. Assuming this hypothesis, we have the exact counter factual for the treatment units and the estimates of our parameter of interest is unbiased.

Hence, we create a sample where treatment and control groups are balanced using matching techniques. Specifically, we are going to use Mahalanobis matching when matching schools with same characteristics and propensity score matching when matching cities. With the balanced samples, we perform our difference-in-differences estimation.

On top of controlling for observed characteristics, we are also going to control for pre-treatment outcomes. Again, the parallel trends hypothesis is necessary for the identification of the parameter of interest. In this sense, matching on pre-treatment outcomes would make the parallel trends hypothesis more plausible, since we are going to select controls that have similar trends of the treatment units. We create a sample of treatment and controls groups in which the outcomes trends are similar. Operationally, we match treatment units with controls units that have similar lagged outcomes in reference to the treatment year of the treatment unit. We use this approach for sake robustness, since its use is not consensus in the literature. On the one hand, Ryan et al. (2015) show that using propensity score matching resulted in more accurate point estimates. On the other hand, Chabé-Ferret (2017), Lindner and McConnell (2018) and Daw and Hatfield (2018) demonstrate that using matching in pre-treatment outcomes increases bias of difference-in-differences. Hence, the results obtained by this approach are going to be used as an extra argument for our findings.

These are the empirical strategies that are going to be used in this work. Other estimations that will be done use similar empirical strategies, and their estimation equation and specifications will be explained just before their results.

[5] term coined by Rubin (1990). Sometimes referred as the Conditional Independence Assumption.

4 Results

In this chapter we are going to present the results of our estimations. In the first section, the main results: we analyze the estimates of our parameter of interest, and check if the parallel trends hypothesis are plausible. Next, we test for other effects that the program can have, such as impacts on outcomes of other schools in the same city of the treated one, or impacts on composition of students and teachers in the treated school. Later, we present the results for the estimation of possible channels of transmission. We seek to understand how the program may impact the outcome: what characteristics are relevant for the impact of the program. In the last section, we discuss the results and relate them to the literature.

4.1 Main Results

As we have approached in the section 3.2, we estimate θ — the parameter of average impact of the full time schooling program on educational outcomes. We do this for the two samples: high school sample and middle school sample. The results for high school are exposed in the tables 8, 9 and 10. The results for middle school are exposed in table 11.

For the high school sample, we test for a variety of specifications. Our baseline specification estimates the equation 3.1, where the controls are the variables described in 1. School and year fixed effects are included. In tables 8, 9 and 10, our baseline specification is our column (1). From the baseline specification, we add different specifications, and combine them to support our main results.

First, we include state trends, to capture possible unobserved factors that change over time and are particular to each state.

Next, we use the matched on observable characteristics sample and the matched on pre-treatment trends sample. As discussed in 3.2.2, we perform these procedures to address possibles issues with the identification hypothesis. We implement a Mahalanobis matching[1] to match on control heterogeneity and to match on pre-treatment variables[2].

1 Mahalanobis matching use the Mahalanobis distance to find close observations. This distance is based on the sample distributions, where the closeness of two observations is given by the number of standard deviations (or its multidimensional generalization) that exist between them. We choose this methodology instead of propensity score matching because of the few numbers of treatment units in some years. Despite the fact that most of the school that will be treated can be found in the first year (2003), there are some few treated units that appeared late in the data set. For example, 10 treated schools appeared in the data-set only in 2009. In order to consider these schools, we perform the matching in each year, pairing the treatment units that first appeared in the data base with controls in the same year.

2 for more details, see appendix A

Table 8 – Impact of the full-time schooling program on ENEM test - High school sample

	Panel A: Impact on ENEM Objective test Score								
Dependent Variable:	ENEM Multiple Choice Test Score								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment dummy	0.279***	0.301***	0.326***	0.358***	0.236***	0.287***	0.329***	0.322***	0.341***
	(0.0445)	(0.0420)	(0.0652)	(0.0501)	(0.0308)	(0.0372)	(0.0440)	(0.0410)	(0.0349)
Observations	80,840	86,150	33,336	33,336	34,106	34,106	86,150	33,336	33,336
R-squared	0.094	0.118	0.119	0.152	0.092	0.123	0.093	0.120	0.153
Number of schools	8,140	8,179	2,893	2,893	2,925	2,925	8,179	2,893	2,893
State trends		Yes		Yes		Yes			Yes
Matching on observables			Yes	Yes				Yes	Yes
Matching on pre-treatment outcomes					Yes	Yes			
Treatment trends							Yes	Yes	Yes

	Panel B: Impact on ENEM writing test score								
Dependent Variable:	ENEN writing Test Score								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment dummy	0.426***	0.439***	0.427***	0.450***	0.364***	0.369***	0.425***	0.407***	0.446***
	(0.0578)	(0.0769)	(0.0582)	(0.0669)	(0.0452)	(0.0448)	(0.0489)	(0.0483)	(0.0417)
Observations	80,942	86,252	33,354	33,354	35,914	35,914	86,252	33,354	33,354
R-squared	0.082	0.118	0.136	0.178	0.115	0.150	0.086	0.138	0.179
Number of schools	8,146	8,185	2,893	2,893	3,023	3,023	8,185	2,893	2,893
State trends		Yes		Yes		Yes			Yes
Matching on observables			Yes	Yes				Yes	Yes
Matching on pre-treatment outcomes					Yes	Yes			
Treatment trends							Yes	Yes	Yes

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 3.1. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies. Columns (3) and (4) are coefficients estimated in the matched on observable sample, where each treated school was matched to the 10 nearest control schools in terms of observable characteristics. Columns (5) and (6) are coefficients estimated in the matched on pre-treatment outcomes sample, where each treated school was matched to the 3 nearest control schools in terms of pre-outcome variables. See Appendices A for additional information.

Table 9 – Impact of the full-time schooling program on flow indicators (1) - High school sample

Panel A: Impact on High School Promotion Rate									
Dependent Variable:	Promotion Rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment dummy	0.762***	0.673***	0.761***	0.673***	0.688***	0.608***	0.681***	0.689***	0.646***
	(0.113)	(0.112)	(0.111)	(0.115)	(0.103)	(0.0926)	(0.0917)	(0.0949)	(0.0983)
Observations	62,913	62,913	23,610	23,610	26,515	26,515	62,913	23,610	23,610
R-squared	0.070	0.127	0.120	0.171	0.110	0.171	0.071	0.122	0.173
Number of schools	7,907	7,907	2,840	2,840	3,065	3,065	7,907	2,840	2,840
State trends		Yes		Yes		Yes			Yes
Matching on observables			Yes	Yes				Yes	Yes
Matching on pre-treatment outcomes					Yes	Yes			
Treatment trends							Yes	Yes	Yes

Panel B: Impact on High School Failure Rate									
Dependent Variable:	Failure Rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment dummy	-0.264**	-0.461***	-0.308***	-0.461***	-0.303***	-0.483***	-0.333***	-0.345***	-0.311***
	(0.106)	(0.0849)	(0.101)	(0.0892)	(0.0917)	(0.0810)	(0.0797)	(0.0816)	(0.0739)
Observations	62,913	62,913	23,610	23,610	23,726	23,726	62,913	23,610	23,610
R-squared	0.020	0.062	0.025	0.080	0.029	0.073	0.020	0.026	0.083
Number of schools	7,907	7,907	2,840	2,840	2,732	2,732	7,907	2,840	2,840
State trends		Yes		Yes		Yes			Yes
Matching on observables			Yes	Yes				Yes	Yes
Matching on pre-treatment outcomes					Yes	Yes			
Treatment trends							Yes	Yes	Yes

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 3.1. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies. Columns (3) and (4) are coefficients estimated in the matched on observable sample, where each treated school was matched to the 10 nearest control schools in terms of observable characteristics. Columns (5) and (6) are coefficients estimated in the matched on pre-treatment outcomes sample, where each treated school was matched to the 3 nearest control schools in terms of pre-outcome variables. See Appendices A for additional information.

The last specification is the addition of treatment trends. This specification in particular is included to probe the robustness of the identification hypothesis. If the estimate of the parameter of interest goes to zero, we have evidence that the full-time schooling program correlates with trends specific to treatment status. This would happen if schools were being chosen to the treatment because they were in a low base line but were in a path of improving quality. In this case, it is not possible to disentangle the causal effect from possible counfounders.

Table 10 – Impact of the full-time schooling program on flow indicators (2) - High school sample

	Panel A: Impact on High School Drop-out Rate								
Dependent Variable:	Drop-out Rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment dummy	-0.957***	-0.613***	-0.907***	-0.604***	-0.786***	-0.453***	-0.699***	-0.698***	-0.678***
	(0.131)	(0.122)	(0.115)	(0.124)	(0.115)	(0.0888)	(0.0801)	(0.0846)	(0.103)
Observations	62,913	62,913	23,610	23,610	26,661	26,661	62,913	23,610	23,610
R-squared	0.062	0.124	0.112	0.196	0.093	0.166	0.065	0.118	0.199
Number of schools	7,907	7,907	2,840	2,840	3,095	3,095	7,907	2,840	2,840
State trends		Yes		Yes		Yes			Yes
Matching on observables			Yes	Yes				Yes	Yes
Matching on pre-treatment outcomes					Yes	Yes			
Treatment trends							Yes	Yes	Yes
	Panel B: Impact on High School Grade-age Distortion Rate								
Dependent Variable:	Grade-age Distortion Rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment dummy	-0.681***	-0.686***	-0.681***	-0.624***	-0.623***	-0.601***	-0.390***	-0.398***	-0.384***
	(0.0917)	(0.0856)	(0.0879)	(0.0831)	(0.0953)	(0.0738)	(0.0555)	(0.0572)	(0.0459)
Observations	62,833	62,833	23,589	23,589	24,155	24,155	62,833	23,589	23,589
R-squared	0.114	0.196	0.209	0.295	0.162	0.262	0.127	0.232	0.312
Number of schools	7,895	7,895	2,838	2,838	2,819	2,819	7,895	2,838	2,838
State trends		Yes		Yes		Yes			Yes
Matching on observables			Yes	Yes				Yes	Yes
Matching on pre-treatment outcomes					Yes	Yes			
Treatment trends							Yes	Yes	Yes

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 3.1. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.Columns (3) and (4) are coefficients estimated in the matched on observable sample, where each treated school was matched to the 10 nearest control schools in terms of observable characteristics. Columns (5) and (6) are coefficients estimated in the matched on pre-treatment outcomes sample, where each treated school was matched to the 3 nearest control schools in terms of pre-outcome variables. See Appendices A for additional information.

The results indicate that the full-time schooling program has positive and significant effects in all educational outcomes, in the high school sample, for all the specifications proposed above. The impacts on ENEM multiple choice test score ranges from 0.236 to 0.356; on ENEM writing test score from 0.364 to 0.450; on promotion rates from 0.608 to 0.762; on failure rates from -0.483 to -0.264; on drop-out rates from -0.957 to -0.453; and on grade-age distortion rates from -0.686 to -0.384, all significant at 1% level.

Moreover, in the specifications with treatment trends, the estimates are affected a little, but they kept their positive and significant values. This indicates that the parallel trends, at this point, are reasonable to assume, and that our estimates of the full-time schooling program impact are indeed causal.

Intuitively, what do these results suggests? Considering only the worst case estimates in the high school sample estimation, we have that a student in a participating school scores, in average, 0.236 standard deviations higher in the ENEM multiple choice and 0.364 standard deviations higher in the ENEM writing test score. For the year of 2013, for example, the average and standard deviation for ENEM multiple choice test score are 498.7 and 50.98, for ENEM writing test score are 418.27 and 141.27. Thus, the average student in a participating school average school will have a ENEM score of 502.52348, while in a non-participating school will have a score 482.614 [3]. According to SISU[4], these extra 20 points in ENEM allow the participating schools' students to apply in 30 more courses in public universities across Brazil in comparison to their counterparts in non-participating schools.

When it comes to the results on flow indicators, the conservative estimates are: 0.608 sd for promotion rates, -0.264 sd for failure rates, -0.453 sd for drop out rates, and -0.384 sd for grade-age distortion rates. For instance, in the year 2010, the mean and standard deviations for these variables were, respectively: 84.06% and 13.21 pp for promotion rates; 10.04% and 9.02 pp for failure rates; 5.89% and 8.23 pp for drop-out rates; and 24.23% and 20.08 pp for grade-age distortion. In terms of percentage points, we have that, in average, the participating schools have a promotion rate 8pp higher, a failure rate 2.38 pp lower, a drop-out rate 3.72 pp lower, and a grade-age distortion 7,71 pp lower. A high school had, in average, 332 students, in 2010. Hence, our results suggests that, a participating high school in comparison to a non-participating school, in average, passed on 26 more and failed 7 less students. Also, 12 less students had dropped-out school in the participating schools.

These are promising results. High school drop-out is one of the biggest problems in

[3] Each university chooses the weights of each test in their own accord. FIES, the federal program for student financing, weights for the calculation of the average are 4 for multiple choice test and 1 for writing test

[4] *Sistema de Seleção Unificada*, the system created by the ministry of education where public universities offers vacancies to the students based on their ENEM scores

education of Brazil. Neri et al. (2015) points "lack of interest" of the students as one of the main reasons for dropping-out school. The pedagogic practices of the program may be impacting here: the life project, youth protagonism, and pedagogy of presence are mainly focused to give to the students purpose and meaning in going to school. Looking at the results on drop-out rates, they seem to be working.

For the middle school sample, we have two specifications. In table 11, the baseline specification are all the odd columns. Our extra specification is the estimation of 3.1 in a matched on observed characteristics sample. The controls are the variables described in 1. School and year fixed effects are included.

As in the high school sample, the results indicates that the full-time schooling program has positive and significant outcomes in the middle school sample. The impacts on *Prova Brasil* language test score ranges from 0.402 to 0.472; on *Prova Brasil* mathematics test score from 0.499 to 0.551; on promotion rates from 0.437 to 0.588; on failure rates from -0.365 to -0.438; on drop-out rates from -0.149 to -0.256; and on grade-age distortion rates from -0.343 to -0.394, all significant at 1% level, with exception to the impact on drop-out rates in the baseline specification, whose estimate is significant only at 10%.

The estimates for the test scores are larger than the analogous variables for the ENEM test scores, but, for the flow indicators, they are smaller. Again, taking into account the conservative estimates, we have that the impact on *Prova Brasil* language scores is 0.402 sd, and on *Prova Brasil* mathematics scores is 0.499 sd. These are large estimates comparing to other Brazilian programs that impact on *Prova Brasil* (Almeida et al. (2016) found negative impacts for math, Oliveira and Terra (2016) found no impact on both test scores, Oliveira (2010) found a impact of 0.2 sd on test scores). But what does these estimates in standard deviations mean? In the year 2013, the average and standard deviations were, respectively, 246 and 18.16302 for language, and 252 and 19.15024 for mathematics. Then, the program increase, in average, 7,3 points of the mathematics test score, and 9.5 points of the language test score.

Looking at the flow indicators, our conservative estimates are 0.437 sd for promotion rates, -0.365 sd for failure rates, -0.149 sd for drop-out rates, and -0.343 sd for grade-age distortion rates. In 2013, the average and standard deviations were, respectively: 88.13% and 9.25 pp for promotion rates;9.180126% and 7.936676 pp for failure rates; 2.687798% and 3.184741 pp for drop-out rates; and 24.9836% and 15.61 pp for grade-age distortion. That means that participating schools, in comparison to non-participating schools, passed-on 4 more students, failed 2 less students, and number of students that are not in the ideal grade according to their age decreased in 5.

Table 11 – Impact of the full-time schooling program on educational outcomes - Middle school sample

Dependent Variable:	Prova Brasil Language Test Score	Prova Brasil Language Test Score	Prova Brasil Mathematics Test Score	Prova Brasil Mathematics Test Score	Promotion Rates	Promotion Rates
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	0.402***	0.472***	0.499***	0.551***	0.437***	0.588***
	(0.111)	(0.0996)	(0.125)	(0.109)	(0.108)	(0.107)
Matching on observed characteristics		yes		yes		yes
Observations	55,196	6,643	55,196	6,643	49,247	5,698
R-squared	0.046	0.062	0.065	0.099	0.023	0.077
Number of Schools	14,020	1,420	14,020	1,420	13,873	1,408

Dependent Variable:	Failure Rates	Failure Rates	Drop-out Rate	Drop-out Rate	Grade-age Distorsion Rate	Grade-age Distorsion Rate
	(7)	(8)	(9)	(10)	(11)	(12)
Treatment dummy	-0.365***	-0.438***	-0.149*	-0.256***	-0.343***	-0.394***
	(0.107)	(0.0797)	(0.0823)	(0.0492)	(0.110)	(0.119)
matching on observed characteristics		yes		yes		yes
Observations	55,166	6,642	55,166	6,642	55,156	6,642
R-squared	0.021	0.070	0.044	0.129	0.061	0.166
Number of Schools	14,003	1,420	14,003	1,420	14,002	1,420

Source: School Census, *Prova Brasil*, INEP, IBGE, MEC/FNDE, Administrative Data. *Note:* Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 3.1. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies. Columns (3) and (4) are coefficients estimated in the matched on observables sample. Columns (5) and (6) are coefficients estimated in the matched on pre-treatment outcomes sample. See Appendices A for additional information.

These results are positive, but they do not seem that impressive in comparison to high school results. That is no coincidence: as we approached in chapter 1, problems of drop-out and grade-age distortion are relatively solved. The marginal improvement in terms of percentage points will be a improvement for less and less students.

We then proceed to visually test the robustness of the parallel trends assumption with leads and lags (graphically). We estimate the equation 3.2, and we plot the estimates of the leads and lags through years. As discussed in the subsection 3.2.1, assuming that the parallel trends hypothesis is true, we have that $\alpha_j = 0$ for $\forall j < 0$, that is, all leads coefficient are equal to zero. Figures 13(a), 13(b), 13(c), 13(a), 13(b) and 13(c) are plots with the point estimates of leads and lags and their respective intervals of confidence, for each outcome of interest. Ideally, we expect that the estimates of leads and lags to be zero, or, at least, to be insignificant.

The results indicate that we should not reject the hypothesis of leads equal to zero. To put it another way, we should not reject the hypothesis of parallel trends. This is another argument, besides the inclusion of treatment trends, for the reasonability of the parallel trends hypothesis.

The leads estimates are not significant, but the point estimates are not exactly zero. Roth (2018) points to the possibility of bias in the difference-in-differences estimator in the case of non-zero pre-trends and non rejection of the hypothesis of parallel trends.

To account for the possible problem of non-zero pre-trend point estimates, we performed a match on pre-treatment outcomes, as discussed in the beginning of this section. We select control units that are similar to the treatment units in terms of their lagged outcomes. Thus, in this sample, we are comparing treatment with control units that have similar pre-outcome trends, enforcing this way the parallel trends hypothesis. If the estimated parameter of interest, in this sample, is similar to the estimated parameter of interest in the whole sample, we have evidence that it is reasonable to assume parallel trends in the whole sample.

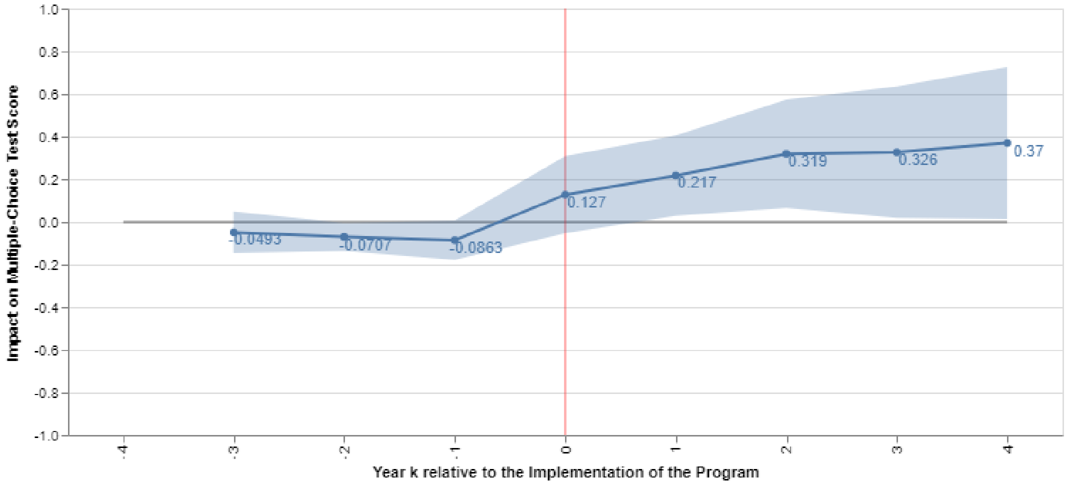
We use the Mahalanobis' metric matching method to attain this goal. Specifically, for each outcome, we create a sample where treated and control units have similar trends for this outcome. For example, for ENEM multiple test scores, we match each treated to control schools whose 2 past lagged ENEM multiple test scores are the closest to the treated unit's 2 past lagged ENEM multiple test scores.

As we have seen, the estimates of the parameter of interest, in the matched on pre-treatment trends, are significant and positive. However, under this specification, does the visual test on the leads and lags improve? In figures 13(a), 13(b) and 13(c) we have the leads and lags graphs of the estimation of 3.2 in the matched on pre-treatment outcomes sample.

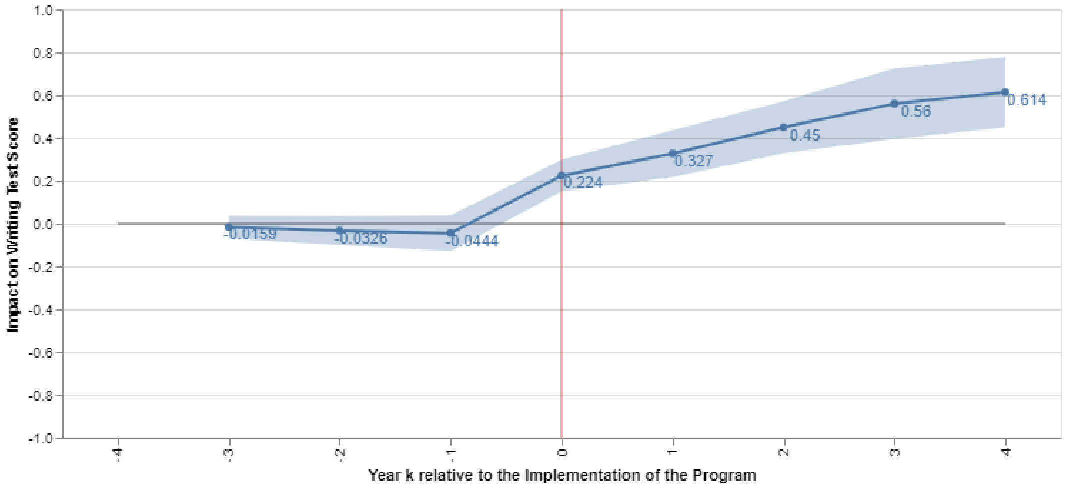
The results of the leads and lags estimation, in the matched on pre-treatment outcomes sample, indicates that we can't reject the hypothesis that the parameters associated to the leads are equal to zero. In like manner, the parallel trends hypothesis are not rejected. Together with the estimation with treatment trends, this suggests that our estimates of the full-time schooling impact on educational are indeed causal.

In summary, we have reasons to believe that the full-time schooling program impacts significantly and in the desired direction[5] on educational outcomes. We tested for a variety of specifications, in different samples, and the results seem robust.

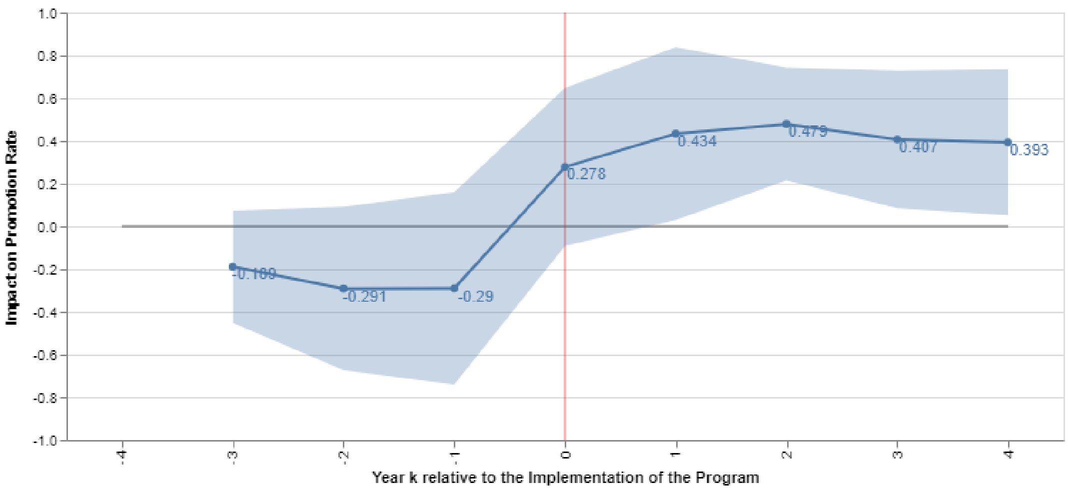
[5] Since for test scores and for promotion rates, positive impacts are socially desirable, and for failure, drop-out and grade-age distortion rates, negative impacts are socially desirable, whenever the directions of impacts follow this pattern, we will denote them as "desired direction". If the impacts are the opposite, that is, negative impact for test scores and promotion rates, and positive impacts to failure, drop-out and grade-age distortion rats, we will address the impacts that follow this patter as "undesired direction".



((a)) Leads and lags for ENEM Multiple choice test Score

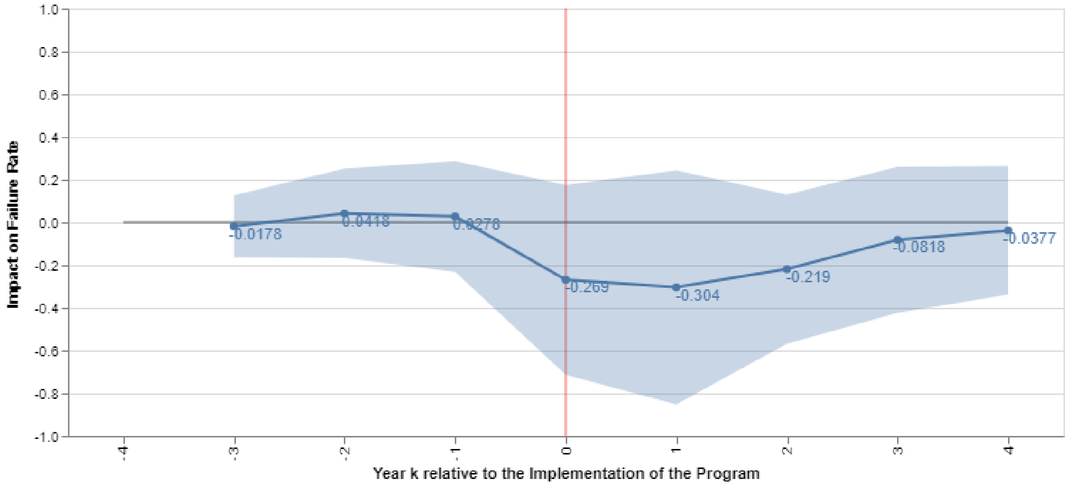


((b)) Leads and Lags for ENEM Writing Test Score

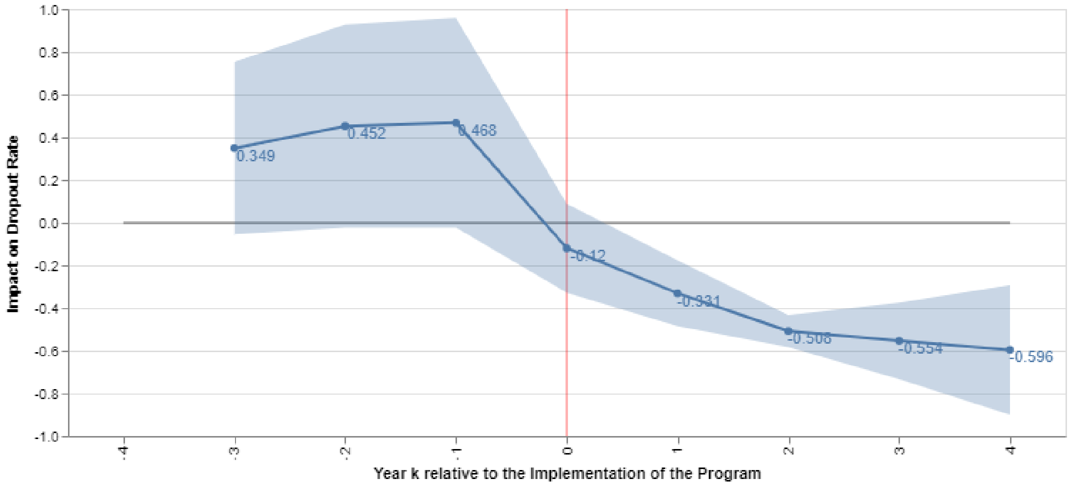


((c)) Leads and lags for promotion rate

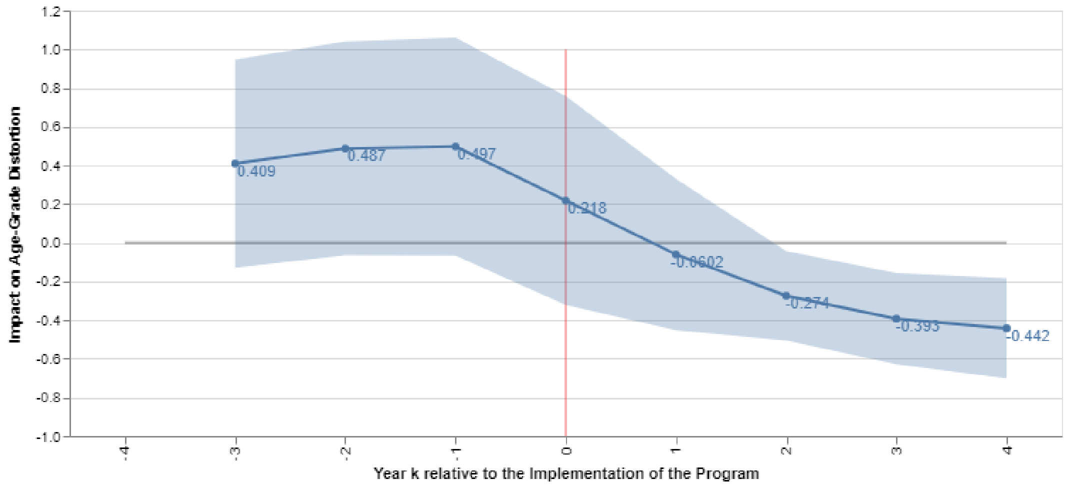
Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The three panels display regression coefficients and associated 95% confidence intervals for the leads and lags of equation 3.2 over years, where k= 0 denotes the year that the school entered the full-time schooling program.



((a)) Leads and lags for failure rate

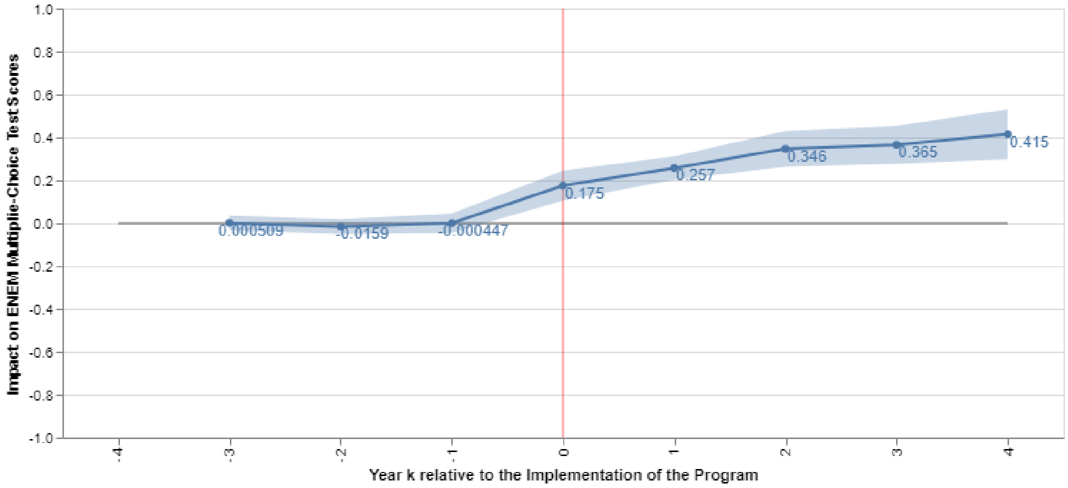


((b)) Leads and lags for drop-out rate

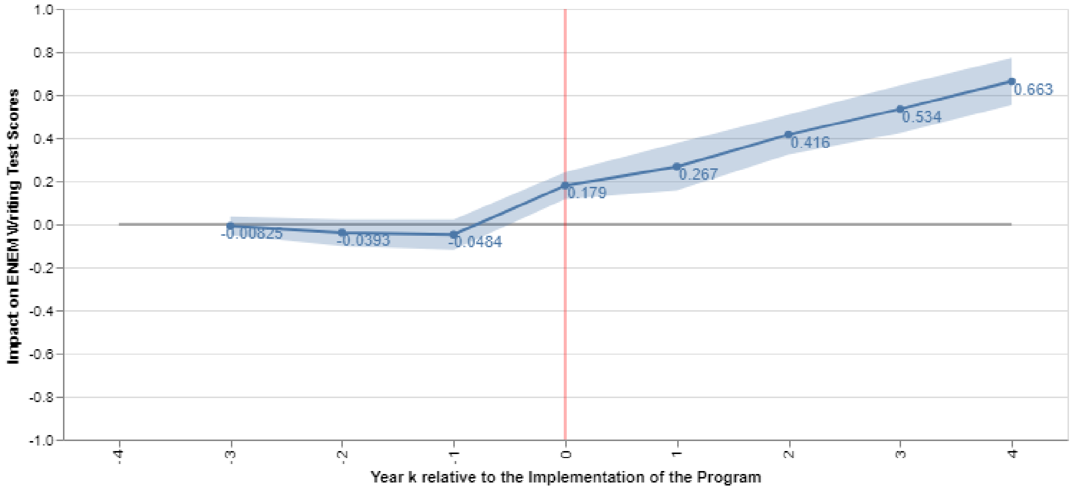


((c)) Leads and lags for grade-age distortion Rate

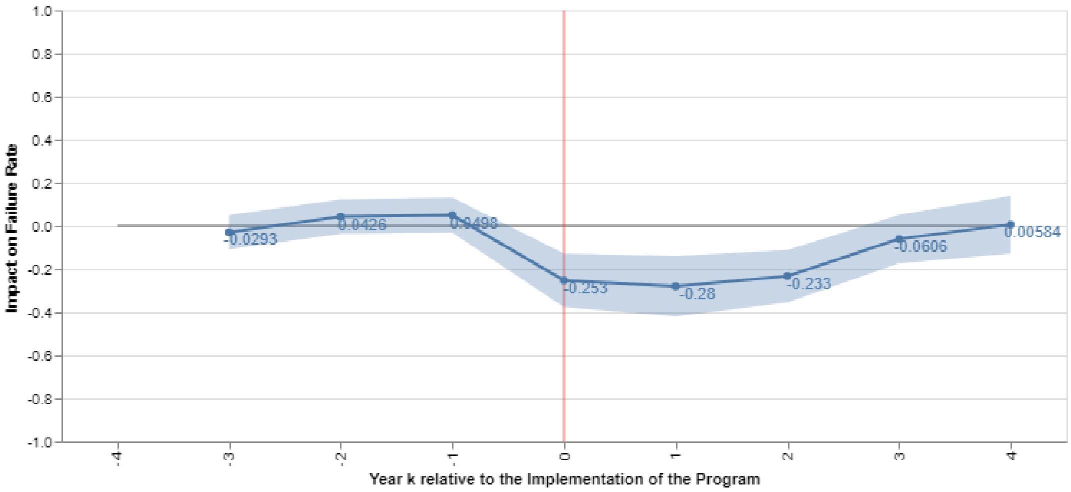
Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The three panels display regression coefficients and associated 95% confidence intervals for the leads and lags of equation 3.2 over years, where k= 0 denotes the year that the school entered the full-time schooling program.



((a)) Leads and lags for ENEM multiple choice test score - Matched on pre-treatment outcomes sample



((b)) Leads and lags for ENEM writing test score matched on pre-treatment outcomes sample



((c)) Leads and lags for failure rate matched on pre-treatment outcomes sample

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The three panels display regression coefficients and associated 95% confidence intervals for the leads and lags of equation 3.2 over years, for the sample of schools matched on pre-treatment outcomes. k= 0 denotes the year that the school entered the full-time schooling program.

4.2 Other Effects

The full-time schooling program can have effects other than the impacts we estimated on test scores and flow indicators. For example, the arrival of the program in one school in the city could make a reallocation of students in the schools. Students (or parents) that want to invest more in their education, knowing that the program is coming to a given school in their city, could potentially want to change to this school. Since students that invest more have better grades, we could have a self-selection effect. The same is valid for teachers: since the program provides better management practices and better pay with bonus incentives, better teachers could be choosing to give classes in the participating schools. The impacts that we estimate in section 4.1 could reflect these reallocation, and not the program itself. The self-selection problem is common in the charter schools literature: if the assignment of students in charter schools is not random, the analysis of charter schools could be hindered, since students are likely to differ simply because they have chosen to attend charter schools, ultimately creating bias when comparing charter schools with non-charter schools (Zimmer and Buddin (2006)).

4.2.1 Spillover

Given the potential problem of self-selection, we are going to check if the program had impact on the other schools educational outcomes. Does the coming of the program in the city cause other schools to worsen its quality?

For this, we use a new sample, a panel of cities, where the treatment units are cities that had school that received treatment. We average the outcomes and controls excluding from the calculation the treated schools. This way, treated cities' variables values are averages that take into account only values from non-treated schools in this same city. We regress the following equation:

$$y_{ct} = \beta_c X_{ct} + \theta_c treatment_{ct} + C_c + T_t + \varepsilon_{ct} \tag{4.1}$$

where y_{ct} is a averaged outcome of interest (test scores for both ENEM and *Prova Brasil* and flow indicators (promotion, failure, dropout and grade-age distortion rates)), X_{ct} is a vector of controls for the city c in the year t, $treatment_{ct}$ is a dummy that indicates that city c in year t is being treated, that is, if the program is implemented in this year or if it was implemented before, in this school. C_c are the school fixed effects, T_t are year fixed effects, and ε_{ct} are unobserved terms.

In the end, we want to see if the coming of the full-time schooling program in the city has impact on the average outcome of others schools of the city. The table 12

reports the results. We test for two specifications: one considering a sample of all cities, and another one matching cities on observable characteristics, using propensity score matching[6].

Table 12 – Spillover effect - Impact of the full-time schooling program on city average educational outcomes - High school sample

Dependent Variable:	ENEM multiple choice test score	ENEM multiple choice test score	ENEM writing test score	ENEM writing test score	Promotion Rate	Promotion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	-0.0818***	-0.0701***	-0.0429	-0.0342	-0.0100	-0.0382
(City has school that implemented the program)	(0.0165)	(0.0161)	(0.0256)	(0.0233)	(0.0531)	(0.0466)
Observations	18,088	7,229	18,088	7,229	12,549	5,423
R-squared	0.067	0.123	0.094	0.127	0.029	0.032
Number of cities	1,430	607	1,430	607	1,430	607
Matching on observable characteristics		Yes		Yes		Yes

Dependent Variable:	Failure Rate	Failure Rate	Drop-out Rate	Drop-out Rate	Grade-age Distortion	Grade-age Distortion
	(7)	(8)	(9)	(10)	(11)	(12)
Treatment dummy	0.117**	0.125***	-0.0629	-0.0378	0.194***	0.149***
(City has school that implemented the program)	(0.0437)	(0.0386)	(0.0981)	(0.0809)	(0.0377)	(0.0251)
Observations	12,549	5,423	12,549	5,423	12,546	5,422
R-squared	0.034	0.053	0.035	0.059	0.073	0.136
Number of cities	1,430	607	1,430	607	1,430	607
Matching on observable characteristics		Yes		Yes		Yes

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given city has a school that is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 4.1. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. All regressions include city fixed effects and year dummies. Columns (2), (4), (6), (8), (10) and (12) are coefficients estimated in the matched on observable sample, where treated schools are paired to the 3 closest schools in terms of observable characteristics. See Appendix A for additional information.

The results are reported in table 12. Interestingly, we have significant impacts on some outcomes, but on others we do not. For instance, for both specifications, we have that the impact of the program on the following average outcomes are significant and in the undesired direction: on the ENEM multiple choice test score, the impacts range from -0.0818 to -0.0701; on the failure rate, the impacts range from 0.117 to 0.125; and on the grade-age distortion, the impacts range from 0.149 to 0.194. The impact on other educational outcomes are insignificant, but in the undesired direction also.

These results suggests that the coming of the program in the city caused non-participating schools to worsen in these outcomes. This can be interpreted in some ways.

6 See A for more information.

Interested students may have changed from the non-participating school to the participating schools, because they or their parents seek for a better quality school, or better teachers from other schools got offers from the participating schools.

However, these findings goes against our conclusion in the section 4.1, where we found positive and significant impacts of the full-time schooling program. Still, assuming that there are some self-selection effect in place, they are not responsible for the evolution of outcomes in the participating schools. The size of the spillover effect is dwarfed in comparison to the full-time schooling program impact on the outcomes of the participating schools: for each significant outcome, the size of the spillover effect is at least half of the size of the program's impact, in the baseline case.

4.2.2 Composition Effects

Similarly to the previous subsection, we are going to test for self-selection effects. Here, we check if the program altered the characteristics of the participating schools — composition effects. We go back to the sample of high schools, and we regress the following equation:

$$z_{it} = \beta_z X_{it} + \theta_z treatment_{it} + S_i + T_t + \varepsilon_{it} \tag{4.2}$$

where z_{it} is a characteristics of the school, X_{it} is a vector of controls for school i in the year t, $treatment_{it}$ is a dummy that indicates that school i in year t is being treated, that is, if the program is implemented in this year or if it was implemented before, in this school. S_i are the school fixed effects, T_t are year fixed effects, and ε_{it} are unobserved terms.

Our parameter of interest, in this case, is θ_z. It is going to capture the effect of the full-time schooling on a set of schools characteristics, including students characteristics, teachers characteristics, and physical characteristics of the school. In the case of total absence of self-selection bias, we would expect that the estimate of θ_z to be not different from zero, regardless of the dependent variable. The program should impact only the outcomes (and some other changes that the full-time schooling program requires the school to do).

Table 13 reports results on averaged students characteristics obtained through ENEM; table 14 reports results on school physical characteristics; and table 15 reports results on number of students and teachers.

The results indicate that the program affects some of students characteristics, as shown in table 13. The impacts on variables "Lives with more than six people", and "Mother has college degree", albeit small, suggest that the type of student in treatment schools are different in characteristics that are usually correlated with years of study, income, and better family structure. Since it is less plausible that the program may be changing these

Table 13 – Impact of the full-time schooling program on average students characteristics - High school sample

Dependent Variable:	Concluded High School in the same year as doing ENEM	Lives with more than six people	Father has college degree	Mother has college degree	Family Income is 5 minimum wages or higher	Looked or is looking for a job
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	14.16***	-0.0227***	0.00101	0.00933**	0.0213***	-0.179***
	(3.271)	(0.00483)	(0.00198)	(0.00415)	(0.00276)	(0.0125)
Observations	87,086	86,488	86,397	86,456	86,488	86,373
R-squared	0.255	0.187	0.033	0.089	0.311	0.645
Number of Schools	8,193	8,190	8,189	8,189	8,190	8,186

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 4.2. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

variables, we assume that students from the other school with these characteristics are coming from non-participating schools. In fact, it seems that some composition effects are taking place, but, again, they are small significant effects (-0.0227, 0.00933, 0.0213 for "Lives with more than six people", "Mother has college degree" and "Family income is 5 minimum wages or higher", respectively). The large and significant estimate for "Looked or is looking for a job" and "Concluded High School in the same year as doing ENEM" are in line with the expected impacts of the full-time schooling program: students in the participating schools are expected to do ENEM when they finish school and to focus in their school life as a consequence of the program's pedagogic objectives.

How do these results translates in number of students? In 2010, the average number of students in a high school was 332. Each of the variables "Lives with more than six people", "Mother has college degree", "Family income is 5 minimum wages or higher" and "Looked or is looking for a job" indicates the proportion of students whose answers for these questions are yes in the school. Assuming that the average characteristic of a high school is the same as the average characteristic of high school student that does ENEM, we have that the program: (i) decreased the number of students that live with more than six people in 7; (ii) increased the number of students whose mother has college degree in 3; (iii) increased the number of students whose family income is six minimum wages or higher in 7; (iv) decreased the number of students who are looking or looked for job in 60.

Indeed, it's reasonable to assume that a change of composition that results in 10 to 20 more students with the profile rich family with highly educated mother does not impact in the average ENEM test score of the whole school (with 320 students) in 20 points. However, the change in the proportion of students who are looking or looked for a job is

much bigger than the other changes: in average, participating schools have 60 students less have this characteristic. This does not seem a result of self-selection effect, where students change from non-participating to participating schools, otherwise, we would see a change of similar magnitude in the other variables . The program's pedagogic practices may be responsible for this: the life project, youth protagonism, and the pedagogy of presence seems to be giving meaning and purpose in school for the students, impacting in a positive manner their choices in life. Thus, the large impact on the variable "Looked or is looking for a job" corroborates with the impacts on the drop-out rates.

Table 14 – Impact of the full-time schooling program on school characteristics - High school sample

Dependent Variable:	Rural Area	Building	Principal's office	Teacher's room	Library	Internet Access	Waste Collection
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment dummy	-0.000772	0.00967**	0.00857	0.00713	0.122***	-0.00963	0.0262***
	(0.00182)	(0.00361)	(0.0122)	(0.0104)	(0.0309)	(0.0220)	(0.00713)
Observations	87,035	87,007	87,035	87,034	87,035	86,294	87,035
R-squared	0.001	0.002	0.007	0.005	0.262	0.088	0.007
Number of Schools	8,191	8,190	8,191	8,191	8,191	8,186	8,191
Dependent Variable:	**Electricity**	**Water**	**Sewage**	**Computer Laboratory**	**Science Laboratory**	**Sports Court**	**Mais Educação Federal Program**
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treatment dummy	0.000369	0.00422	-0.00988	0.0587***	0.175***	0.0232	0.0573***
	(0.000386)	(0.00457)	(0.0144)	(0.0176)	(0.0240)	(0.0374)	(0.0176)
Observations	87,035	87,035	87,034	87,033	87,033	87,033	87,035
R-squared	0.001	0.003	0.008	0.521	0.053	0.218	0.209
Number of Schools	8,191	8,191	8,191	8,191	8,191	8,191	8,191

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 4.2. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

Looking at the impacts on school characteristics (table 14), we have some significant results. The effects on "Building", "Library", "Computer laboratory", "Science laboratory" are significant and, to some extent, expected. The selection of the schools for the participation is arbitrary, and secretary of education choose schools that, among other operational criteria, have infrastructure that could receive the program. This means that schools that have laboratories, library, school cafeteria, building have higher probability to be chosen. Also, in the program's early stage, the implementation of the program included construction of these facilities. This could also be driving the estimate up. On the other hand, the significance of the parameter of interest on "Waste collection" and "*Mais Educação* Federal Program" are somewhat unexpected. Waste collection does not seem to

be a necessary operational criteria for the implementation of the program. Therefore, the estimate indicates that the full-time schooling program made some participating schools to have waste collection. Since the program implements new management practices, and there are selection and removal of school principals, it makes sense that better management would make the school adopt some better practices in relation to waste. In the case of the impact on the variable "*Mais Educação* Federal Program", we have that participating schools in the full-time schooling program have incentives to seek for the federal program, which is voluntary, and finances costs associated with the extension of school day. The school has some autonomy on how to spend resources. Given that the full-time schooling program for itself is an extension of school day, one would expect principals from the treated schools to seek for the federal program to extra funding for the school.

One could argue that schools with *Mais Educação* federal program are being selected to participate the full-time schooling program, and that the impacts found in the section 4.1 are in fact results from the Federal Program. This seems not to be the case. The estimate of impact on participation of the federal program is 0.0573. That is, schools participating in the full-time schooling program participate, in average, 5.73% more in the federal program. It is not plausible that the impact seen in 4.1 are consequence of this variation in participation of *Mais Educação* federal program. Additionally, as we have seen in the introduction, estimations of the impact of the federal program are small or even negative.

Table 15 reports the impact of the full-time schooling program on number of students and teachers. The results indicate that the full-time program impacts negatively and significantly on the number of students and teachers. It appears that participating schools enroll less students and hire less teachers, probably indicating some aspect of the management practices. However,it does not seem that there are some kind of selection of better teachers: the impact on the proportion of teachers with college degree is close to zero and insignificant.

The negative impact of the program on number of teachers and on number of teachers with college, and the insignificance of the impact on the proportion of college teachers, suggests that there is no selection effect by the teachers side. It doesn't seem that more educated teachers are leaving non-participating schools to give classes in participating schools. Actually, it seems that less teachers are being employed.

Overall, spillover and composition effects indicate that some self-selection effect is taking place, but they are small in comparison to the impact of the program in educational outcomes. Non-participating schools in the same city of participating schools perform worse in the ENEM test and in the flow indicators, but the size of this effect pales in comparison to the size of the full-time schooling program's impact on educational outcomes. This suggests that some good students and teachers may have changed from non-participating

Table 15 – Impact of the full-time schooling program on number of students and teachers - High school sample

	Panel A: Impact on Number of Students		Panel B: Impact on number of Teachers			
Dependent Variable:	Number of Students	Number of Female Students	Number of Teachers	Number of Teachers with High School Degree	Number of Teachers with College Degree	Proportion of Teachers with College Degree
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	-117.0***	-55.55***	-22.72***	-1.462***	-21.20***	0.000229
	(27.80)	(14.70)	(4.894)	(0.436)	(4.947)	(0.00725)
Observations	86,324	86,265	86,265	86,265	86,265	86,265
R-squared	0.283	0.249	0.613	0.100	0.608	0.082
Number of Schools	8,203	8,185	8,185	8,185	8,185	8,185

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 4.2. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

schools to participating schools. The impact on composition reinforces this idea: the program alters the composition of students, shifting to richer students, with better family structure; alters the composition of school characteristics, but in expected variables; and alters the number of students and teachers. However, all these effects are not large enough to have account for the full-time schooling program.

4.3 Possible Channels of Transmission

Impacts of full-time schooling and extension of school day are ambiguous in the literature. In its review, Alfaro et al. (2015) analyzed 19 studies that measure the effect of longer school days. The results are mixed, some significant and negative, such as Llambí (2013) with impacts of -0.29 standard deviations, other are significant and positive, such as Hincapie (2016), with impacts of 0.05 to 0.16 standard deviations.

As we approached in the introduction, the literature of extension of school time is not conclusive, and other characteristics of the extension of school time programs may be relevant to the impact on education. One of this work's objective is to cast some light in this ambiguity. Not only estimate the full-time schooling program's impacts, but also understand which are the channels through which more time impact educational impacts.

In the previous sections, we found results that indicate that the full-time schooling

program has indeed a significant impact on educational outcomes. Moreover, we found that the there are some spillover and selection effects, but they are dwarfed in terms of magnitude when faced to the values of the impact's estimate. Note that the full-time schooling program is a program that implements not only the extension of school time, but also pedagogic and management practices.

Hence, in this section, we are going to analyze these other characteristics, or the channels of transmission, as we are going to call them. We use some variation present in the implementation of the implementation of the program to identify some of the channels; compare some estimations from sub-samples from early and late stages, whose implementations differs in some fundamental characteristics; and analyze the overflow effect, which can indicate that some characteristics of the program are indeed more relevant than others.

4.3.1 Mediators

In our data set, we have some variation of how the program has been implemented in the schools. The following questions were made to the team that helped the implementation of the program in each school:

1. Was state governor engaged with the implementation of the program?
2. Was the secretary of education engaged with the implementation of the program?
3. Was there a dedicated team in Secretary of education for the implementation of the program?
4. Did changes in the legal framework take place?
5. Did all necessary changes in the legal framework take place, by the prescribed deadline?
6. Did selection and removal of principals process take place?
7. Did selection and removal of teachers process take place?
8. Was the life project implemented in the curriculum?

With these questions and respective answers, we created dummy variables that indicate 1 when the answer is yes and 0 otherwise. We denote each of these dummies as a "mediator". Mediator 1 is the dummy corresponding to the question 1, mediator 2 to the question 2, and so on. We interact them with the treatment dummy, and include them in the equation

Table 16 – Estimating the mediators' effect

Dependent Variable:	ENEM multiple choice test score	ENEM writing test score	Promotion Rate	Failure Rate	Drop-out Rate	Grade-age Distortion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	0.667***	0.742***	0.336***	-0.299***	-0.236*	-0.440***
	(0.191)	(0.229)	(0.104)	(0.0920)	(0.121)	(0.0952)
Mediator 1 * treatment dummy	-	-0.280***	-	-	-	-
		(0.0961)				
Mediator 2 * treatment dummy	-0.491*	0.205	1.169***	-0.705***	-1.123***	-0.0371
	(0.274)	(0.366)	(0.304)	(0.184)	(0.301)	(0.182)
Mediator 3 * treatment dummy	0.238***	-	0.551***	-0.574***	-0.265	0.125*
	(0.0442)		(0.170)	(0.0646)	(0.232)	(0.0623)
Mediator 4 * treatment dummy	-0.335*	-0.187	0.567***	-0.134*	-0.763***	-0.552***
	(0.189)	(0.236)	(0.0685)	(0.0697)	(0.0761)	(0.101)
Mediator 5 * treatment dummy	-0.0185	-0.379	-1.902***	1.014***	1.979***	0.632***
	(0.269)	(0.311)	(0.331)	(0.145)	(0.372)	(0.180)
Mediator 6 * treatment dummy	-0.109	0.283	1.284***	-0.846***	-1.187***	-0.599***
	(0.267)	(0.332)	(0.325)	(0.144)	(0.369)	(0.178)
Mediator 7 * treatment dummy	0.527*	0.212	-1.347***	1.223***	0.840	-0.0423
	(0.312)	(0.398)	(0.481)	(0.236)	(0.540)	(0.285)
Mediator 8 * treatment dummy	-	-	-	-	-	-
Observations	86,150	86,252	62,913	62,913	62,913	62,833
R-squared	0.094	0.088	0.071	0.023	0.067	0.121
Number of Schools	8,179	8,185	7,907	7,907	7,907	7,895

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The table reports coefficients from the difference-in-differences estimation from equation 3.1 with the inclusion of interactions between treatment dummy and the mediators. All variables are at city level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

3.1. Our goal here is to estimate the impact of each one the specific characteristics in the implementation of the program.

Table 16 reports the results. Most of the estimates of the mediators interaction are not significant when the dependent variable is a test score variable (with the exception to the mediator 1 interaction — "Was state governor engaged with the implementation of the program?" — whose estimate is surprisingly significant and negative). In the case of flow indicators, the scenario is different. All mediators interactions have a significant impact on promotion rate. Some are positive, reaching 1 standard deviation of impact (Mediator 2 — "Was the secretary of education engaged with the implementation of the program?" and Mediator 6 — " Did selection and removal of principals process take place?"), but other are negative, reaching almost -2 standard deviations of impact (Mediator 5 — "Did all necessary changes in the legal framework take place, by the prescribed deadline?").

Taking a closer look, we can notice that there is something wrong. Besides the fact that the estimates are unreasonably large, mediator 1 and mediator 8 interactions were dropped from the estimation. In table 28 we have the correlation matrix of the mediators, which shows that there are mediators that are high correlated with others. Despite the

fact that we have answers for the questions about implementations, there are not many variance between them, leading to the problem of multicolinearity.

To deal with this problem, we aggregate the mediators that share same characteristics in groups. We test for two kind of dummy aggregations. First, we create dummies that assume 1 if all the related mediators also assume 1, and 0 otherwise. Next, we create dummies that assume 1 if any of the related mediators assumes 1, and 0 otherwise. We do this aggregations with mediators 1 to 7, since each one of them has another mediators that they are high correlated with.

Table 17 – Estimating the mediators' effect - Aggregating dummies - All

Dependent Variable:	ENEM multiple choice test score	ENEM writing test score	Promotion Rate	Failure Rate	Drop-out Rate	Grade-age Distortion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	0.666***	0.743***	0.335***	-0.302***	-0.232*	-0.439***
	(0.192)	(0.228)	(0.104)	(0.0924)	(0.121)	(0.0951)
All political dummy * treatment dummy	-0.301***	-0.415***	-0.148	-0.143	0.411**	0.671***
	(0.0577)	(0.0611)	(0.145)	(0.163)	(0.169)	(0.0620)
All legal framework dummy * treatment dummy	0.187	-0.151	-0.637**	0.444**	0.546*	-0.464***
	(0.142)	(0.164)	(0.245)	(0.179)	(0.281)	(0.120)
All selection dummy * treatment dummy	-0.0717	0.420	1.107***	-0.323***	-1.476***	-0.680***
	(0.242)	(0.271)	(0.127)	(0.0800)	(0.130)	(0.0810)
Mediator 8 * treatment dummy	-0.491*	-0.0745	1.170***	-0.695***	-1.136***	-0.0409
	(0.275)	(0.322)	(0.303)	(0.185)	(0.300)	(0.182)
Observations	86,150	86,252	62,913	62,913	62,913	62,833
R-squared	0.094	0.088	0.071	0.023	0.066	0.121
Number of Schools	8,179	8,185	7,907	7,907	7,907	7,895

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: We create a dummy called "All political dummy", which assumes 1 when the mediators 1, 2 and 3 ("Was state governor engaged with the implementation of the program?", " Was the secretary of education engaged with the implementation of the program?", and "Was there a dedicated team in Secretary of education for the implementation of the program?") assumes 1 at the same time, and 0 otherwise; a dummy called "All legal framework dummy", which assumes 1 when the mediators 4 and 5 ("Did changes in the legal framework take place?" and "Did all necessary changes in the legal framework take place, by the prescribed deadline?") assumes 1 at the same time, and 0 otherwise; a dummy called "All selection dummy", which assumes 1 when the mediators 6 and 7 ("Did selection and removal of principals process take place?" and "Did selection and removal of teachers process take place?") assumes 1 at the same time, and 0 otherwise. The table reports coefficients from the difference-in-differences estimation from equation 3.1 with the inclusion of interactions between treatment dummy and the aggregated dummies plus the interaction of mediator 8. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

We them proceed to regress 3.1 with the inclusion of these aggregated dummies. One specification with the "all" dummmies and mediator 8 and another specification with the "any" dummies and mediator 8. Tables 17 and 18 reports the results of these estimations.

Table 18 – Estimating the mediators' effect - Aggregating dummies - Any

Dependent Variable:	ENEM multiple choice test score	ENEM writing test score	Promotion Rate	Failure Rate	Drop-out Rate	Grade-age Distortion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	0.444***	0.725***	0.429***	-0.226***	-0.467**	-0.575***
	(0.108)	(0.142)	(0.148)	(0.0741)	(0.182)	(0.138)
Any political dummy * treatment dummy	-0.288***	-0.456***	-0.0779	-0.180	0.330*	0.568***
	(0.0567)	(0.0890)	(0.150)	(0.146)	(0.184)	(0.108)
Any legal framework dummy * treatment dummy	-	-	-	-	-	-
Any selection dummy * treatment dummy	0.327***	0.332***	0.290	0.0685	-0.575*	-0.880***
	(0.0602)	(0.0821)	(0.192)	(0.157)	(0.298)	(0.130)
Mediator 8 * treatment dummy	-0.0947	-0.167	0.370**	-0.291***	-0.274	-0.266*
	(0.0838)	(0.102)	(0.170)	(0.0930)	(0.302)	(0.146)
Observations	86,150	86,252	62,913	62,913	62,913	62,833
R-squared	0.094	0.088	0.071	0.023	0.066	0.121
Number of Schools	8,179	8,185	7,907	7,907	7,907	7,895

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: We create a dummy called "Any political dummy", which assumes 1 when any of the mediators 1, 2 and 3 ("Was state governor engaged with the implementation of the program?", " Was the secretary of education engaged with the implementation of the program?", and "Was there a dedicated team in Secretary of education for the implementation of the program?") assumes 1, and 0 otherwise; a dummy called "Any legal framework dummy", which assumes 1 when any of the mediators 4 and 5 ("Did changes in the legal framework take place?" and "Did all necessary changes in the legal framework take place, by the prescribed deadline?") assumes 1, and 0 otherwise; a dummy called "Any selection dummy", which assumes 1 when any of the mediators 6 and 7 ("Did selection and removal of principals process take place?" and "Did selection and removal of teachers process take place?") assumes 1, and 0 otherwise. The table reports coefficients from the difference-in-differences estimation from equation 3.1 with the inclusion of interactions between treatment dummy and the aggregated dummies plus the interaction of mediators 8. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

The estimates of the interactions of the "All political dummy " and the "Any political dummy" are all in the undesired direction, with some of them being significant, and others not. For instance, the impact of the "All political dummy" interaction on ENEM writing test score is -0.415 standard deviations . This suggests that the political engagement is somewhat bad to the implementation of the program: schools that had some political actor engaged in the implementation seems to perform worse than the their counterparts. The estimates of the interactions of the "All legal framework dummy" are also in the undesired direction as well, but the "Any legal framework dummy" was dropped from the estimation.

On the other hand, the estimates of the interactions related to the selection process of teachers and principals and the mediators 8, related to the implementation of life project in the curriculum, are in general in the desired direction. For example, the impacts related to the "All selection dummy" interaction on Promotion rate is 1.107, on failure rate is -0.323, on drop-out rate is -1.476, and on grade-age distortion is -0.680. The impacts related to the mediators 8 — "Was the life project was implemented in the curriculum?" — on promotion rate and on failure rate are 0.370 and -0.291, respectively, in the specification with the "any" dummies.

Overall, the estimations using this aggregated dummies show a direction, a sign of the channels' impact. However, they are still unreliable, due to the high correlation between the mediators and to the multicolinearity problem. To probe for more robustness of these results, we proceed to a new approach for multicolinearity: principal component analysis. The idea of this approach is to reduce the dimensionality of the mediators, extracting common variance between them and transforming into a new set of orthogonal variables, the principal components. These components will retain most of the variation present in all mediators and best explain them. For more information, go to appendix B.

After applying this method on the mediators 1 to 7 (the ones that have high correlation between them and possibly are the source of the multicolinearity problem), we obtain a new set of two components, that accounts for most of the variance of the mediators: a component more correlated to the political mediators — mediators 1 to 5 —, and a component more correlated to the selection mediators — mediators 6 and 7[7]. These two components are going to be included in the regression, along with the mediator 8, all of them interacted with the treatment dummy.

Table 19 reports the results of the regression with the components interactions. The results follow the same pattern observed in the estimation with aggregation of dummies: for the component associated to the political mediators, the correspondent estimates are not significant when the dependent variables are test scores, but are significant when the dependent variables are flow indicators, though they go in the undesired direction.

[7] for details of how the components were chosen and the interpretation, see B

For the component associated to the selection component and for the mediator 8, the estimates of their interaction with the treatment are not all significant, but they all go in the desired direction. For instance, the impact of the interaction of selection component and the treatment dummy on ENEM multiple choice test score is 0.133; on ENEM writing test score is 0.0906; on promotion rate is 0.161; on failure rate is -0.237; on drop-out rate is -0.251; and on grade-age distortion is -0.332. Despite the point estimates, this estimation is trick. How can we interpret a change in the components and its impact in the outcome? It is a question that make innocuous the size of the estimates. Nevertheless, the sign of the estimates are consistent with the estimations with aggregation dummies.

In summary, in this subsection, we seek to estimate some channels of transmission using the mediators. We found the problem of multicolinearity and, to address this, we aggregated the mediators that are high correlated to each other and used a principal component analysis. The point estimates are unreliable, but the sign and significance of the estimates suggests that the political mediators, associated with the engagement of political actors in the implementation of the program in the school, worsen the impacts of the program. On the other hand, they suggest also that the selection mediators, associated with the implementation of selection and removal processes, and the mediator 8, that indicate that the life project was implemented in the curriculum, improve the impacts of the program.

Table 19 – Estimating the mediators' effect - Components

Dependent Variable:	ENEM multiple choice test score	ENEM writing test score	Promotion Rate	Failure Rate	Drop-out Rate	Grade-age Distortion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment dummy	0.321***	0.548***	0.411***	-0.246***	-0.414**	-0.496***
	(0.115)	(0.144)	(0.144)	(0.0811)	(0.179)	(0.112)
Political Component * treatment dummy	0.0546	-0.00854	0.0968***	-0.0559***	-0.106*	-0.138***
	(0.0348)	(0.0411)	(0.0330)	(0.0183)	(0.0618)	(0.0177)
Selection Component * treatment dummy	0.133***	0.0906*	0.161**	-0.0237	-0.251**	-0.332***
	(0.0464)	(0.0531)	(0.0597)	(0.0394)	(0.100)	(0.0256)
Mediator 8 * treatment dummy	0.0834	-0.258	0.731***	-0.562***	-0.601	-0.611***
	(0.230)	(0.264)	(0.225)	(0.135)	(0.387)	(0.135)
Observations	86,150	86,252	62,913	62,913	62,913	62,833
R-squared	0.094	0.088	0.071	0.023	0.067	0.121
Number of Schools	8,179	8,185	7,907	7,907	7,907	7,895

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: Treatment dummy indicates that a given schools is being treated in a given year. The table reports coefficients from the difference-in-differences estimation from equation 3.1 with the inclusion of interactions between treatment dummy and the components plus the interaction of mediator 8. All variables are at school level. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies.

4.3.2 Difference in implementations

The full-time schooling program suffered several changes through years. In the years 2004 to 2007 – the early stage — the program was implemented only in one state. In this stage, the Institute that created and idealized the program had a important role in the implementation in schools. Together with the managing body in the secretary of education, they helped to train employees, principals, teachers and implement all new methodologies in schools. Also, they had a important role in making the necessary changes in the legal framework for the program to function adequately. The Institute role was essential in the success in implementing all the practices they idealized.

After, in the late stage (2008-2015), the program changed its *modus operandi*. It expanded to other states. The Institute stepped out of any operational role in the secretaries, and acted only as a consultant group for the states governments. The role that in the early stage belonged to the Institute, in the late stage, was assumed by the secretary of education. From 2008 on, When the partnership is firmed, the Institute goes to the state government, trains its employees of the secretary of education, and transfer the "technology" to implement the program in schools. One big difference was that, in the early stage, the presence of the Institute reinforced the implementation of all practices and theories idealized for the program. In the late stage, the secretaries received recommendations of implementation, but they could ultimately choose not to follow.

Since we have this variance in implementation, we are going to estimate the impact of the program in the two stages and compares the size of the point estimates. We want to capture the difference in impact of a rigorous implementation, characterized by the presence of the institute in the process, and other less rigorous implementations. In the estimation, we regress the main equation 3.1 in the following samples: in a sample containing only schools of the first state, in the years of the early stage (2004 - 2007), which we call "Sample early stage"; in a sample containing schools from the first state, but from 2004 to 2015, which we denote as "Sample first state all years"; and in a sample containing schools from all states from the late years.

Table 20 reports the impacts on ENEM test scores[8], estimated in the different samples. All the estimates are significant at the 1% level. Comparing the point estimates, the larger ones are from the estimations made in the sample "early stage" (0.530 for ENEM multiple choice test score and 0.661 for ENEM writing test score). They are almost twice the size of the point estimates relative to the sample "all state late years" (0.247 and 0.358 for ENEM multiple choice test score and for ENEM writing test score). When looking at the estimations made in the sample "first state all years", they are almost 0.2 standard

[8] Flow indicators were not available until 2008, when the early stage has ended

Table 20 – Different implementations

Dependent Variable:	ENEM multiple choice test score				ENEM writing test score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment dummy	0.279***	0.315***	0.530***	0.247***	0.426***	0.385***	0.661***	0.358***
	(0.0445)	(0.0338)	(0.0713)	(0.0365)	(0.0578)	(0.0491)	(0.0641)	(0.0515)
Observations	80,840	8,729	4,896	56,426	80,942	8,753	4,905	56,528
R-squared	0.094	0.106	0.054	0.067	0.082	0.192	0.170	0.073
Number of codigo_escola	8,140	949	863	7,838	8,146	951	865	7,844
main sample	Yes				Yes			
Sample early stage			Yes				Yes	
Sample first state all years		Yes				Yes		
Sample all states late years				Yes				Yes

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The table reports coefficients from the difference-in-differences estimation from equation 3.1. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies. Each column correspond to the estimation made in a given sample, for the ENEM test score outcomes. The estimated coefficients are displayed side-by-side to comparison between the impacts of each sample.

deviations smaller than the correspondent point estimates related to the sample "early stage".

These results suggest that program in the early stage is more effective in bringing a larger impact than the late stage of the program. This is reasonable: the program implementation was rigorous in this stage, since the institute was closer, more present, and had influence in the executing body of the program in secretary of education. Since one of the biggest differences between the implementations in early and late years were characterized by the absence of relative politically costly practices in the late years, the difference in results between different implementations suggests that these same practices are responsible for the larger impacts in the early years. Among these practices, we have the system of incentives and process of selection and removal of teachers. These findings go accordingly to the findings in the previous subsection 4.3.1: process of selection and removal of teachers and principals seem to be a characteristic that make the program impact more.

4.3.3 Overflow effects

In this subsection, we estimate the overflow effects of having a program in one educational stage on outcomes of the other educational stage, that does not have the

program. Here we want to know if some characteristics of the program that make changes in the whole school, not only in the educational stage it was primarily implemented, are drivers of the impacts seen in the section 4.1. For instance, one would expect that the implementation of selection of principals and the adoption of new management practices would impact the whole school. On the other hand, the implementation of the life project in the curriculum and the full-time schooling in one educational stage are not expected to impact the other educational stage. This is what we are trying to capture: are there overflow effects? does the program impact through changes that impacts the whole school or through changes that affect only the targeted educational stage?

In our both samples, we have schools which have both middle and high schools. Additionally, among them, there are those which received the treatment in only one educational stage. Accordingly, we are going to estimate the following equation 4.3:

$$z_{it} = \beta_z X_{it} + \theta_z treatment\_other_{it} + S_i + T_t + \varepsilon_{it} \qquad (4.3)$$

where y_{it} is a averaged outcome of interest of the educational stage a (test scores for both ENEM or *Prova Brasil* and flow indicators (promotion, failure, dropout and age-grade distortion rates)) where $a \in$ {middle school; high school}, X_{it} is a vector of controls for the school i in the year t, $treatment_{it}$ is a dummy that indicates that school i in year t is being treated in the educational stage $b \in$ {middle school; high school}, where $b \neq a$, but it is not being treated in the educational stage a. S_i are the school fixed effects, T_t are year fixed effects, and ε_{it} are unobserved terms.

Table 21 – Overflow effect - High school sample

Dependent Variable:	ENEM multiple choice test score	ENEM writing test score	Promotion Rate	Failure Rate	Drop-out Rate	Grade-age Distortion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
School has program only in the Middle School stage	-0.0101	0.201*	0.749***	-0.965***	-0.0811	-0.199**
	(0.113)	(0.111)	(0.176)	(0.225)	(0.167)	(0.0898)
Observations	80,839	80,941	62,912	62,912	62,912	62,832
R-squared	0.086	0.072	0.049	0.018	0.041	0.062
Number of Schools	8,140	8,146	7,907	7,907	7,907	7,895

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The table reports coefficients from the difference-in-differences estimation from equation 4.3. The dependent variables are educational outcomes for the High School Sample. The treatment dummy assumes 1 when the high school in the school does not have the program but the middle school has, and zero otherwise. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies. Each column correspond to the estimation made in a given sample, for the ENEM test score outcomes.

Tables 21 and 22 report the results for both high school and middle school sample.

The implementation of the program in only one education stage appears to have some impact in the other non-participating educational stage. In the high school sample, the presence of the program only in the middle school has significant impact on high school promotion rate (0.749), high school failure rate (-0.965), and high school grade-age distortion (-0.199). In the middle sample, the impacts of the presence of the program only in high school on middle schools is significant in most of the educational outcomes and goes in the desired direction, with exception to *Prova Brasil* Language Test Scores, which is not significant. Overall, looking at the results of the estimations on both samples, we can see that the largest and most significant impacts are on flow variables, particularly on promotion and failure rates, and the impact on test scores are barely significant.

These difference between impacts on test scores vis-à-vis impacts on flow indicators suggests that the full-time schooling program affects these group of variables through different channels. Programs characteristics that make changes in the whole school, such as goal setting, selection and removal of teachers and principals, evaluation of principals and teachers, and the adoption of system of incentives, tend to impact more flow indicators. Programs characteristics that changes only the target educational stage, such as pedagogic practices and full-time schooling, are not the drivers for flow indicators. Since the program does have impact on test scores, one possible hypothesis is that pedagogic practices are channels of transmission of the impact for test scores.

Table 22 – Overflow effect - Middle school sample

Dependent Variable:	Prova Brasil Language Test Score	Prova Brasil Mathematics Test Score	Promotion Rates	Failure Rates	Drop-out Rate	Drop-out Rate
	(1)	(2)	(3)	(4)	(5)	(6)
School has program only in the High School stage	0.110	0.159**	0.457***	-0.201**	-0.553***	-0.224**
	(0.0812)	(0.0692)	(0.102)	(0.0925)	(0.122)	(0.0837)
Observations	55,196	55,196	49,247	55,166	55,166	55,156
R-squared	0.045	0.062	0.023	0.019	0.044	0.058
Number of Schools	14,020	14,020	13,873	14,003	14,003	14,002

Source: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data.
Note: The table reports coefficients from the difference-in-differences estimation from equation 4.3. The dependent variables are educational outcomes for the Middle School Sample. The treatment dummy assumes 1 when the middle school in the school does not have the program but the High school has, and zero otherwise. ***significance at the 1% level; **significance at the 5% level; *significance at the 10% level. Robust standard errors, clustered at mesoregion level, are reported in parenthesis. Clustering in state, city, and city year levels do not change the significance of the estimates. All regressions include school fixed effects and year dummies. Each column correspond to the estimation made in a given sample, for the ENEM test score outcomes.

4.4 Discussion

All these results suggest that the program indeed has impact on educational outcomes. Considering the program only as a extension of time program, these results contrasts somehow with the results found in other programs and policies in Brazil, like in the program *Escola de Tempo Integral*, whose impact was either small and positive, or insignificant (Aquino (2011)), or like in the *Mais Educação* Federal Program, whose impacts are nonexistent or negative (Almeida et al. (2016), Oliveira and Terra (2016)).

Further analysis shows that some specific characteristics of the program may be the channels through which the program may impact the educational outcomes. Using the mediators, we found the implementation of a process of selection and removal of teachers and principals and the implementation of the life project may be some of the characteristics that improve the impact. On the other hand, the engagement of political entities in the implementation of the program in school may worsen the impact. This is somewhat surprising, since one would expect that the engagement of important political actors would make the implementation of the practices easier. One explanation to this negative impact indicate that political parties are somehow associated with the principal and teachers of the schools whose implementation they are engaged in, and thus, process of selection and removal are not exact applied.

Using differences in the implementations, we found that the early stage's impacts were larger, which suggests that the rigorousness in the implementation — consequence of the presence of the institute in the process — is important for the impact of the program. This reflects the fact that the absence of some characteristics in the late stage, such as funding and hard political practices implementations, such as teacher's bonus incentive and removal of teachers, may be relevant to the impact of the program. This could be related with the charter school literature. Dias and Guedes (2010) denotes this early stage of the program as a Brazilian charter school, since the Institute, as private entity had some influence in the management of the participating schools and contributed with funds to fixed costs of new schools. In the literature, charter schools seems to have a mixed effect: Booker et al. (2007) find that students experience poor test score growth in their initial year in charter schools, with subsequent recovery. However, there is weak evidence that there may be overall gains from charter attendance. In the other hand, Hoxby and Kang (2009) and Dobbie and Jr (2013) find that charter attendance improves test outcomes. Our findings seems to agree with the later.

At last, analyzing overflow effects of the program, we found that the program impacts not only the educational stage they are being implemented in, but also the other educational stage. This suggests that some characteristics of the program that impacts the whole school, such as a management practices — goal setting, system of incentives, selection and removal of teachers and principals, trainings, and planning — improves the

impact of the program on overflow effects. Similarly, the absence of overflow effects on test score and the existence of impact of the program on test score suggests that the program's characteristics that are restricted to one educational stage are the channels through which the program impact test scores.

Overall, from the mediators, different implementations and overflow effects analysis, management aspects such as goal setting, selection and removal of teachers and principals, trainings, teacher's incentive, and teacher and principal evaluation seem to be important to the impact of the program. This could be related with the literature of school management and teacher's incentive: Lavy (2009) found that individual monetary incentives led to significant improvements in test taking rates, conditional pass rates, and mean test scores. Muralidharan and Sundararaman (2011) and Glewwe et al. (2003) finds that providing incentives improves students test score. Crawfurd (2017) finds that school management quality matters for student value-added in Uganda; Di Liberto et al. (2015) found that managerial practices are positively related to student outcomes, and suggest that principal selection and training should be implemented. Bloom et al. (2015) show that high management quality is strongly associated with better educational outcomes. Tavares (2015) concludes that specific management practices such as performance monitoring, targets setting and incentive schemes have significant positive impacts on students math-scores. Our findings contribute in reinforcing these results in the literature: management and teacher's incentive are important for improving educational outcomes.

5 Conclusion

In this work, we estimate the impact of the Full-Time schooling program on educational outcomes, such as test scores and flow variables. The program is characterized by the extension of school time and by the adoption of new pedagogic and management practices. Some of these pedagogic practices are based on the concept such as integral education, pedagogy of presence, youth protagonism, and full-time schooling. The management practices are based on a management instrument for goal setting and achievement, a system of incentives for teachers and school employees, selection of teachers and school managers, and others. In the literature of full-time schooling and extension of school time, the results are ambiguous, and we seek to understand how a full-time schooling program can impact educational outcomes. Therefore, besides the estimation of impacts of the full-time program, we use some variation in the implementation of the program to estimate the possible channels of transmission through which the program may impact the outcomes.

Our results relies on a dynamic difference-in-difference strategy. We test and assume the parallel trends assumption to identify our parameter of interest, the average impact of the program. For sake of robustness, we perform a matching on observed characteristics, to address the possible selection bias (the treatment assignment is not random) and we perform a matching on pre-treatment outcomes, to enforce the parallel trends hypothesis.

Schools that have been selected for the program have a higher average ENEM score (0.279 standard deviations for multiple choice score and 0.426 standard deviations for writing score), a higher *Prova Brasil* score (0.4 to 0.47 standard deviations) higher promotion rates (0.762), lower failure, drop out and grade-age distortion rates (-0.264, -0.957, and -0.681)[1]. In the matched samples, the results does not differ much. Translating these standard deviations: in 2010, average students in a participating school had 30 more courses in public universities to choose, with their scores, vis-à-vis to students in non-participating schools; participating schools passed-on 26 more students, failed 7 less; 12 less students had dropped-out schools.

Additionally, we test for other possible effects of the program. We estimate the impact of the program in school, such as average students' characteristics, and school characteristics, using the same methodology as before. We found that schools that receive the program have more students doing the ENEM test in the year they conclude high school (14.16), a smaller proportion of students that lives with more than six people in

[1] All these results reported in this paragraph are the from the baseline specification

their households (-0.022), a higher proportion of students whose mother concluded college (0.009), a bigger proportion of students whose family income is higher than 5 minimum wages (0.02) and a smaller proportion of students who worked or looked for work (-0.17). Interestingly, there is no impact in the proportion of students whose father concluded college (0.001). Despite the significance of the effects, the magnitude of most of them are small, what indicates that the impact on educational outcomes are not resulting for selecting students with these characteristics. When it comes to school characteristics, the program seems not to have much impact. For many of the schools physical characteristics, there is no significant effect. However, there are some exceptions: schools that received the program have in average more libraries (12.2 p.p.), more waste collection (2 p.p.), and more laboratories (5 p.p. for science laboratories and 17.5 p.p. for computer laboratories). Schools that received the program also participated more in the federal program *Mais Educação*, in average (5p.p.).

Moreover, we estimate the impact of the program on the average outcomes of the city, excluding the treated schools from the calculation. With this, we seek to analyze the impact of the coming the program in a school of the city had in the other schools of the same city. Our results indicate that the program indeed had impact on the average of the whole city. For ENEM, we have a negative impact of 0.08 on multiple choice test score. For flow indicators, we have a positive impact of 0.117 on failure rate and 0.149 on grade-age distortion. This suggests that some selection effect really took place, but the magnitude is smaller than the magnitude of the effect of the program.

Given the significant impact of the program on the outcomes, we them proceed to analyze how the program got these impacts. We seek to understand what are the channels of transmission of the program's impact. First, using dummies that indicate some differences in the implementation of the program, we found that, when the school implemented teacher and school manager selection process, the impact of the program is bigger. This goes in line with the literature of management and teacher's incentive.

Second, we use the difference in implementation between early years (2004-2008) and late years (2009-2015). The early years were characterized by a bigger influence of the private institute in the treated schools management, like a private public partnership, which allowed them to enforce all the program practices. This partnership is sometimes characterized as a charter school experience in Brazil. When comparing estimates from these periods, we find that the impact on schools in the early period is bigger, what suggests that this private public partnership is somehow better for the efficiency of the program. This result adds to the discussion of charter schools in developing countries.

Third, we estimate the impact of having the program in one educational stage on the other educational stage that did not receive the program. That is, in schools that have both educational stages, but only have received the program in one stage, we

estimate the impact on the other educational stage that did not receive it. We seek to estimate possible overflow effects: changes in behaviour in the whole school that were caused by the implementation of the program in only one stage. The overflow effects are not significant for test scores, but significant for flow indicators. This result suggest that some practices of the program, despite being implemented in only one stage, have effect in the whole school. Looking at the design of the program, some practices impact all the teachers and employees of the school, such as the system of incentives, selection process and other management practices. These are probably the channels for the program to the flow indicators. Other program's characteristics, that can be "contained" in only on stage, may be the responsible for the impact on test scores, such as the extension of school day and changes in curriculum.

In short, we have that the full-time schooling program increases test scores and improves the flow indicators of treated schools. There are some selection and composition effects, but they don't seem to be the main cause of the impact on educational outcomes. Furthermore, we find that teachers and school managers selection process is a relevant channel of transmission. We contribute to the full-time schooling and extension of school day literature with a study analyzing the impacts of a full-time schooling program in Brazil, that also can be related the management and school incentives literature.

Bibliography

Abadie, A. (2005). Semiparametric difference-in-differences estimators.

Abbiati, G. and E. C. Meroni (2016). How do students react to longer instruction time? evidence from italy. *Education Economics 24*(6), 592–611.

Agüero, J. M. (2016). Evaluación de impacto de la jornada escolar completa.

Alfaro, P., D. K. Evans, P. Holland, et al. (2015). Extending the school day in latin america and the caribbean. *Policy Research Working Paper 7309*.

Almeida, R., A. Bresolin, B. Borges, K. Mendes, and N. Menezes-Filho (2016). *Assessing the impacts of Mais Educacao on educational outcomes: evidence between 2007 and 2011*. The World Bank.

Anderson, M. and M. B. Walker (2015). Does shortening the school week impact student performance? evidence from the four-day school week. *Education Finance and Policy 10*(3), 314–349.

Aquino, J. M. (2011). *A ampliação da jornada escolar melhora o desempenho acadêmico dos estudantes? Uma avaliação do programa Escola de Tempo Integral da rede pública de São Paulo*. Ph. D. thesis, University of São Paulo.

Autor, D. H. (2003). *Journal of labor economics*, 1–42.

Battistin, E. and E. C. Meroni (2016). Should we increase instruction time in low achieving schools? evidence from southern italy. *Economics of Education Review 55*, 39–56.

Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? results from a natural experiment in chile. *Economics of Education Review 28*(5), 629–640.

Bloom, N., R. Lemos, R. Sadun, and J. Van Reenen (2015). Does management matter in schools? *The Economic Journal 125*(584), 647–674.

Bonilla, L. (2011). Doble jornada escolar y calidad de la educación en colombia.

Booker, K., S. M. Gilpatric, T. Gronberg, and D. Jansen (2007). The impact of charter school attendance on student performance. *Journal of Public Economics 91*(5-6), 849–876.

Burkam, D. T., V. E. Lee, D. D. Ready, J. Honigman, and S. J. Meisels (2006). Full-day versus half-day kindergarten: In which program do children learn more? *American Journal of Education 112*(2), 163–208.

Cannon, J. S., A. Jacknowitz, and G. Painter (2006). Is full better than half? examining the longitudinal effects of full-day kindergarten attendance. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management 25*(2), 299–321.

Cerdan-Infantes, P. and C. Vermeersch (2007). *More time is better: An evaluation of the full time school program in Uruguay.* The World Bank.

Chabé-Ferret, S. (2017). Should we combine difference in differences with conditioning on pre-treatment outcomes? TSE Working Papers 17-824, Toulouse School of Economics (TSE).

Crawfurd, L. (2017). School management and public–private partnerships in uganda. *Journal of African Economies 26*(5), 539–560.

Daw, J. R. and L. A. Hatfield (2018). Matching in difference-in-differences: between a rock and a hard place. *Health services research 53*(1), 4111–4117.

DeCicca, P. (2007). Does full-day kindergarten matter? evidence from the first two years of schooling. *Economics of Education Review 26*(1), 67–82.

Di Liberto, A., F. Schivardi, and G. Sulis (2015). Managerial practices and student performance. *Economic Policy 30*(84), 683–728.

Dias, M. C. N. and P. M. Guedes (2010). O modelo de escola charter: a experiência de pernambuco. Technical report.

Dobbie, W. and R. G. F. Jr (2013). Getting beneath the veil of effective schools: Evidence from new york city. *American Economic Journal: Applied Economics 5*(4), 28–69.

Farbman, D. and C. Kaplan (2005). Time for a change: The promise of extended time schools for promoting student achievement. executive summary. *Massachusetts 2020*.

Gandra, J. M. d. F. (2017). O impacto de educação em tempo integral no desempenho escolar: uma avaliação do programa mais educação. Master's thesis, Federal University of Viçoca.

Gibbs, C. (2014). Experimental evidence on early intervention: The impact of full-day kindergarten. *Frank Batten School of Leadership and Public Policy Working Paper 4*.

Glewwe, P., N. Ilias, and M. Kremer (2003). Teacher incentives.

Hincapie, D. (2016). Do longer school days improve student achievement? evidence from colombia. *IDB Working paper series*.

Hoxby, Caroline M., S. M. and J. Kang (2009). How new york city's charter schools affect achievement. Technical report.

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

James-Burdumy, S., M. Dynarski, M. Moore, J. Deke, W. Mansfield, C. Pistorino, and E. Warner (2005). When schools stay open late: The national evaluation of the 21st century community learning centers program. final report. *US Department of Education*.

Junior, R. J. L. H. (2011). O Ginásio Pernambucano e o Procentro: os desafios para transformar um projeto piloto em política de rede. Master's thesis, Federal University of Pernambuco.

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review 99*(5), 1979–2011.

Lindner, S. and K. J. McConnell (2018). Difference-in-differences and matching on outcomes: a tale of two unobservables. *Health Services and Outcomes Research Methodology*, 1–18.

Link, C. R. and J. G. Mulligan (1986). The merits of a longer school day. *Economics of Education Review 5*(4), 373–381.

Llambí, M. C. (2013). El efecto causal de la política de tiempo completo sobre los resultados educativos en la enseñanza media: aplicación de cuatro métodos no experimentales e identificación de posibles sesgos.

Magalhães, M. (2008). *Pernambuco cria, Experimenta e Aprova.*

Meroni, E. C. and G. Abbiati (2016). How do students react to longer instruction time? evidence from italy. *Education Economics 24*(6), 592–611.

Meyer, E. and C. Van Klaveren (2013). The effectiveness of extended day programs: Evidence from a randomized field experiment in the netherlands. *Economics of Education Review 36*, 1–11.

Muralidharan, K. and V. Sundararaman (2011). Teacher performance pay: Experimental evidence from india. *Journal of political Economy 119*(1), 39–77.

Neri, M. et al. (2015). Motivos da evasão escolar.

OECD (2015). Education at a glance 2015: Oecd indicators. Technical report.

OECD (2017). Education at a glance 2017: Oecd indicators. Technical report, OECD.

Oliveira, J. M. d. (2010). Custo-efetividade de políticas de redução do tamanho da classe e ampliação da jornada escolar: uma aplicação de estimadores de matching.

Oliveira, L. F. B. d. and R. Terra (2016, August). Impacto do Programa Mais Educação em indicadores educacionais. Working Papers Portuguese 147, International Policy Centre for Inclusive Growth.

Pereira, G. C. (2011). Uma avaliação de impacto do programa mais educação no ensino fundamental.

Robin, K. B., E. Frede, and W. S. Barnett (2007). *Is more better?: The effects of full-day vs. half-day preschool on early school achievement.* National Institute for Early Education Research, Rutgers.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Roth, J. (2018). Should we adjust for the test for pre-trends in difference-in-difference designs? *arXiv preprint arXiv:1804.01208*.

Rubin, D. B. (1980). Bias reduction using mahalanobis' metric matching. *Biometrics 36*(2), 293–298.

Rubin, D. B. (1990). Formal mode of statistical inference for causal effects. *Journal of statistical planning and inference 25*(3), 279–292.

Ryan, A. M., J. F. B. Jr, and J. B. Dimick (2015). Why we should not be indifferent to specification choices for difference-in-differences. *50*(4), 1211–1235.

Tavares, P. A. (2015). The impact of school management practices on educational performance: Evidence from public schools in são paulo. *Economics of Education Review 48*, 1–15.

Votruba-Drzal, E., C. P. Li-Grining, and C. Maldonado-Carreño (2008). A developmental perspective on full-versus part-day kindergarten and children's academic trajectories through fifth grade. *Child Development 79*(4), 957–978.

World Bank (2018). World development report 2018: Learning to realize education's promise. Technical report.

Zimmer, R. and R. Buddin (2006). Charter school performance in two large urban districts. *Journal of Urban Economics 60*(2), 307–326.

Zvoch, K., R. E. Reynolds, and R. P. Parker (2008). Full-day kindergarten and student literacy growth: Does a lengthened school day make a difference? *Early Childhood Research Quarterly 23*(1), 94–107.

Appendix

APPENDIX A – Matching On Observed Characteristics

Matching on observed characteristics is a procedure that we use to balance our control and treatment units, as approached in the section 3.2.2. Our main objective is the plausiblity of the uncounfoundedness hypothesis. With this, we seek to give more robustness to our estimations of our parameter of interest.

We use this procedure two times: (i) to match treated to control schools and (ii) to match treated to control cities. In the first, we use the mahalanobis' metric matching; in the second, we use the propensity score matching. In the following sections discuss these methods and their implementations in our study.

A.1 Matching Schools Using the Mahalanobis' Metric

We use the Mahalanobis's metric to match schools. First approached by Rubin (1980), this matching technique takes into account correlations across covariates and leads to matches that are invariant to affine transformations of the covariates Rubin (1990). We choose to use this method over propensity score matching because we match treated (actually, we are going to match schools that will be treated in the future against schools that will never be treated. For sake of simplicity, we are going to address the first group as treated and the second as control, in this section) to control units in the first year that the treated unit appears in our database. Most of our treated schools are present in the first year of our database, 2003. However, there are some treated schools which appear late in the database. In order not to lost information about them, we match each treated school with control units in the first year they appear in the database. Since the number of new treated schools in a given year could be low, we can't estimate a propensity score properly, and mahalanobis' metric comes in hand.

Formally, the mahalanobis' distance between observation 1 and observation 2, in terms of their covariates x_1 and x_2, is given by:

$$d_M(x_1, x_2) = (x_1 - x_2)S_{tc}^{-1}(x_1 - x_2)^T$$

where

$$S_{tc} = \left(\frac{N_c.\hat{\Sigma}_c + N_t.\hat{\Sigma}_t}{N_c + N_t}\right)$$

with N_c and N_t being the number of controls and treated, respectively and $\hat{\Sigma}_c$ $\hat{\Sigma}_t$ being the sample covariance matrices, for control and treatment groups, respectively.

The procedure consists in the following steps: we fix a year. For each school, we see if they belong to the treatment group, that is, if they are being or are going to be treated. If that is the case, we check if this is the first year in the database. If yes, we calculate the mahalanobis' distance between this school with each the controls, in terms of the covariates[1]. We select the 10 nearest control schools. For this treated school and the 10 control schools, we keep all the observations of the following years. We do this procedure for all years, and, in the end, we have a sample of treatment units and controls units that are near to the treatment ones in terms of the mahalanobis' distance.

In tables 23 and 24 we have the sample balancing after matching, in terms of average characteristics, by treatment status, for high school and middle school samples. note that the first year in the high school sample is 2003, and the first year in the middle school sample is 2007. The matching seems to have balanced the samples reasonably. Comparing to the sample balancing before the matching, in tables 6 and 7, we can see that in the matched samples, the mean differences of variables are less significant, in general. It is reasonable to not reject the uncounfoundedness assumption in the matching samples.

[1] the covariates used here are displayed in tables 23 and 24

Table 23 – Sample Balancing after matching - Average school characteristics, in 2003, by treatment status - High school sample

year = 2003	Nonparticipating Schools		Schools that will participate in the Full-time Schooling Program		Difference (1)-(3)	
Variable	Mean	sd	Mean	sd	Mean Diff	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
ENEM writing test score	-0.45	0.61	-0.47	0.54	0.02	(0.72)
ENEM multiple choice test score	-0.46	0.49	-0.50	0.44	0.04	(1.47)
Concluded High School in the same year as doing ENEM	91.13	80.54	99.79	84.55	-8.65	(-1.79)
Lives with more than six people	0.11	0.09	0.13	0.10	-0.02***	(-4.33)
Father has college degree	0.03	0.05	0.04	0.04	-0.00	(-1.48)
Mother has college degree	0.03	0.06	0.04	0.04	-0.00	(-1.81)
Family Income is 5 minimum wages or higher	0.14	0.13	0.11	0.13	0.03***	(3.59)
Looked or is looking for a job	0.84	0.12	0.80	0.11	0.03***	(5.22)
Rural Area	0.02	0.15	0.01	0.09	0.01*	(2.53)
Building	1.00	0.04	0.99	0.09	0.01	(1.33)
Principal's office	0.95	0.22	0.93	0.25	0.02	(1.13)
Teacher's room	0.95	0.21	0.95	0.22	0.01	(0.49)
Library	0.80	0.40	0.86	0.35	-0.06**	(-2.96)
Internet Access	0.76	0.43	0.82	0.38	-0.07**	(-2.88)
Waste Collection	0.98	0.14	0.96	0.20	0.02*	(2.01)
Electricity	1.00	0.00	1.00	0.00	0.00	(.)
Water	0.96	0.19	0.96	0.20	0.00	(0.37)
Sewage	0.75	0.44	0.78	0.42	-0.03	(-1.34)
Computer Laboratory	0.59	0.49	0.58	0.49	0.01	(0.46)
Science Laboratory	0.38	0.48	0.45	0.50	-0.07*	(-2.49)
Sports Court	0.81	0.39	0.77	0.42	0.04	(1.84)
Number of Students in High School	594.69	412.15	653.57	374.17	-58.89**	(-2.67)
Number of Female Students in High School	311.84	220.26	351.15	205.21	-39.30**	(-3.27)
GDP per Capita	22717.78	21294.04	17023.75	18012.12	5694.02***	(5.34)
City Population	1635359.70	3308173.40	719602.92	2033332.91	915756.78***	(7.04)
Mais Educação Federal Program	0.53	0.50	0.40	0.49	0.13***	(4.64)
Number of Teachers in High School	24.61	15.38	27.03	12.75	-2.42**	(-3.19)
Number of Teachers with High School Degree in High School	1.89	5.57	1.13	4.02	0.76**	(3.10)
Number of Teachers with College Degree in High School	22.71	15.86	25.89	13.25	-3.19***	(-4.05)
Proportion of High School Teachers with College Degree	0.91	0.22	0.96	0.16	-0.05***	(-4.68)
Observations	2129		351		2480	

Table 24 – Sample Balancing after matching - Average school characteristics, in 2007, by treatment status - Middle school Sample

year = 2007 Variable	Nonparticipating Schools Mean	 sd	Schools Participating that will participate in the Full-time Schooling Program Mean	 sd	Difference (1)-(3) Mean Diff	 t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
Prova Brasil Language Test Score	0.30	0.76	0.39	0.78	-0.09	(-1.39)
Prova Brasil Mathematics Test Score	0.22	0.72	0.28	0.73	-0.06	(-1.07)
Mother has college degree	0.06	0.05	0.07	0.05	-0.01**	(-2.74)
Father has college degree	0.06	0.05	0.08	0.05	-0.01**	(-3.26)
Rural Area	0.00	0.05	0.00	0.00	0.00	(1.73)
Building	1.00	0.03	1.00	0.00	-0.00	(-1.00)
Principal's office	0.96	0.20	0.96	0.20	-0.00	(-0.01)
Teacher's room	0.97	0.18	0.99	0.11	-0.02*	(-2.05)
Library	0.89	0.31	0.92	0.27	-0.03	(-1.19)
Internet Access	0.94	0.24	0.98	0.13	-0.04***	(-3.38)
Waste Collection	1.00	0.06	1.00	0.00	-0.00*	(-2.00)
Electricity	1.00	0.00	1.00	0.00	0.00	(.)
Water	0.99	0.10	0.99	0.08	-0.00	(-0.47)
Sewage	0.91	0.28	0.96	0.19	-0.05**	(-3.02)
Computer Laboratory	0.87	0.34	0.93	0.26	-0.06**	(-2.77)
Science Laboratory	0.26	0.44	0.40	0.49	-0.15***	(-3.66)
Sports Court	0.89	0.32	0.93	0.26	-0.04	(-1.91)
Number of Class Rooms in use	12.28	3.82	13.57	4.61	-1.29***	(-3.44)
GDP per Capita	33892.11	17998.98	36174.21	18445.79	-2282.10	(-1.50)
City Population	2077380.46	3652512.89	2400297.94	3756273.29	-322917.48	(-1.04)
Mais Educação Federal Program	0.50	0.50	0.63	0.48	-0.13**	(-3.22)
Number of Classes in Middle School	12.92	4.82	12.76	5.54	0.16	(0.35)
Number of Students in Middle School	444.94	195.07	437.19	225.71	7.75	(0.42)
Number of Female Students in Middle School	217.96	96.12	212.13	111.28	5.83	(0.64)
Number of White Students in Middle School	178.83	114.26	180.99	107.34	-2.17	(-0.24)
Number of Teachers in Middle School	115.13	40.73	119.70	41.00	-4.58	(-1.35)
Number of Teachers in High School	69.76	70.27	60.99	65.08	8.76	(1.61)
Number of Teachers with College Degree in High School	111.91	41.41	116.58	42.13	-4.66	(-1.34)
Propotion of Middle School Teachers with College Degree	0.97	0.09	0.97	0.09	-0.00	(-0.52)
Observations	1205		166		1371	

A.2 Matching Cities Using the Propensity Score

Different from the previous section, here we use propensity score to match the cities. We use this method here because we do not have the same issue as we have in the cases of the schools. All the treated cities are present in all years of the sample. Propensity score is defined as the probability of receiving the treatment given the observable covariates Rosenbaum and Rubin (1983). The advantage of this method is that the adjustment for the scalar propensity score is sufficient to remove bias due to all observable covariates, since the balacing score property of the propensity score implies that, conditional on the propenisty score, the assignment is uncounfounded (see Chapter 12 of Rubin (1990) for more information). This way, when we construct a sample where treatment and control units have similar propensity scores, we are construcing balanced sample and we are seeking for the plausability of the uncounfoundedness hypothesis.

Our procedure consist in: (i) estimating the propensity score for each school in the initial year, since we don't have its real value; (ii) selecting 3 controls for each treated city whose estimated propensity scores are nearest the treated city's estimated propensity score; (iii) we them keep their observations in the following years. This way, we construct a sample where treated and control units are similar in terms of estimated propensity score.

Table 25 – Descriptive Statistics - City level - High school sample

Variable	Nobs	Mean	sd
City Population	18,284	59,540	364,242
GDP per Capita	18,284	20,981	27,956
ENEM writing test score	18,277	-0.417	0.396
ENEM multiple choice test score	18,284	-0.303	0.499
Promotion Rate	12,620	-0.0543	0.638
Failure Rate	12,620	-0.137	0.646
Drop-out Rate	12,620	0.254	0.706
Grade-age Distortion Rate	12,617	0.246	0.757
Concluded High School in the same year as doing ENEM	18,284	60.44	46.23
Lives with more than six people	18,234	0.0758	0.0801
Father has college degree	18,232	0.0359	0.0430
Mother has college degree	18,234	0.0706	0.0629
Family Income is 5 minimum wages or higher	18,233	0.0550	0.0702
Looked or is looking for a job	18,227	0.623	0.237
Number of Students in High School	18,240	386.4	254.1
Number of Female Students in High School	18,239	204.6	138.5
Rural Area	18,284	0.0523	0.141
Building	18,282	0.994	0.0539
Principal's office	18,282	0.941	0.182
Teacher's room	18,282	0.945	0.172
Computer Laboratory	18,282	0.783	0.364
Science Laboratory	18,282	0.348	0.387
Sports Court	18,282	0.736	0.369
Internet Access	18,200	0.890	0.269
Waste Collection	18,282	0.978	0.109
Electricity	18,282	0.999	0.0234
Water	18,282	0.956	0.143
Sewage	18,282	0.686	0.409
Mais Educação Federal Program	18,284	0.471	0.424
Number of Teachers in High School	18,284	79.22	61.90
Number of Teachers with High School Degree in High School	18,284	3.149	8.092
Number of Teachers with College Degree in High School	18,284	76.04	59.83
Proportion of High School Teachers with College Degree	18,245	0.950	0.112
Number of Schools in the city	18,284	5.416	24.43

Note: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data

Table 26 – Sample balancing before matching - Average city characteristics, by treatment status - High school sample

Variable	Nonparticipating Cities		Cities that will participate in the Full-time Schooling Program		Difference (1)-(3)	
	Mean	sd	Mean	sd	Mean Diff	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
ENEM writing test score	-0.40	0.43	-0.46	0.37	0.07**	(2.71)
ENEM multiple choice test score	-0.38	0.39	-0.50	0.33	0.13***	(5.94)
Concluded High School in the same year as doing ENEM	58.60	38.30	67.73	44.14	-9.13***	(-3.46)
Lives with more than six people	0.08	0.08	0.16	0.11	-0.08***	(-12.93)
Father has college degree	0.03	0.04	0.03	0.03	0.01***	(3.57)
Mother has college degree	0.06	0.05	0.05	0.04	0.01**	(2.70)
Family Income is 5 minimum wages or higher	0.14	0.11	0.08	0.10	0.06***	(8.87)
Looked or is looking for a job	0.79	0.13	0.77	0.11	0.02**	(2.69)
Rural Area	0.03	0.12	0.06	0.13	-0.03***	(-3.41)
Building	1.00	0.05	0.99	0.07	0.01*	(2.04)
Principal's office	0.95	0.19	0.89	0.25	0.06***	(4.23)
Teacher's room	0.95	0.19	0.91	0.21	0.04**	(2.86)
Library	0.55	0.46	0.66	0.38	-0.11***	(-4.39)
Internet Access	0.75	0.40	0.83	0.31	-0.08***	(-3.92)
Waste Collection	0.98	0.12	0.94	0.19	0.04***	(4.01)
Electricity	1.00	0.03	1.00	0.03	0.00	(0.32)
Water	0.97	0.14	0.93	0.18	0.03**	(2.96)
Sewage	0.69	0.44	0.64	0.39	0.05	(1.95)
Computer Laboratory	0.01	0.08	0.01	0.07	-0.00	(-0.72)
Science Laboratory	0.15	0.31	0.21	0.29	-0.05**	(-2.91)
Sports Court	0.31	0.42	0.16	0.25	0.15***	(8.00)
Number of Students in High School	410.75	247.02	552.71	273.73	-141.96***	(-8.58)
Number of Female Students in High School	215.30	133.23	310.00	153.92	-94.70***	(-10.27)
GDP per Capita	17295.33	20672.45	11976.20	15655.74	5319.13***	(5.06)
City Population	36020.90	200930.96	119325.50	608143.27	-83304.60*	(-2.51)
Mais Educação Federal Program	0.43	0.45	0.47	0.38	-0.04	(-1.68)
Number of Teachers in High School	21.80	9.25	22.47	8.45	-0.67	(-1.25)
Number of Teachers with High School Degree in High School	2.03	3.88	1.08	2.03	0.95***	(5.94)
Number of Teachers with College Degree in High School	19.77	9.65	21.37	8.38	-1.60**	(-2.98)
Proportion of High School Teachers with College Degree	0.89	0.20	0.95	0.09	-0.06***	(-7.23)
Observations	1070		348		1418	

Note: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data

Table 27 – Sample Balancing after matching - Average city characteristics, by treatment status - High school sample

Variable	Nonparticipating Cities		Cities that will participate in the Full-time Schooling Program		Difference (1)-(3)	
	Mean	sd	Mean	sd	Mean Diff	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
ENEM writing test score	-0.37	0.41	-0.45	0.36	0.08*	(2.56)
ENEM multiple choice test score	-0.35	0.34	-0.49	0.32	0.14***	(5.11)
Concluded High School in the same year as doing ENEM	73.09	43.62	75.81	45.87	-2.72	(-0.74)
Lives with more than six people	0.11	0.10	0.15	0.09	-0.04***	(-5.18)
Father has college degree	0.04	0.04	0.03	0.03	0.01**	(3.17)
Mother has college degree	0.06	0.05	0.05	0.04	0.01	(1.63)
Family Income is 5 minimum wages or higher	0.14	0.10	0.10	0.10	0.04***	(4.85)
Looked or is looking for a job	0.79	0.11	0.77	0.11	0.02**	(2.65)
Building	1.00	0.02	0.99	0.08	0.01*	(2.27)
Principal's office	0.93	0.20	0.90	0.22	0.03*	(2.02)
Teacher's room	0.94	0.20	0.93	0.16	0.01	(0.80)
Computer Laboratory	0.02	0.12	0.01	0.08	0.00	(0.36)
Science Laboratory	0.16	0.30	0.21	0.27	-0.05*	(-2.21)
Sports Court	0.25	0.38	0.18	0.25	0.07**	(2.61)
Internet Access	0.86	0.30	0.85	0.28	0.01	(0.32)
Rural Area	0.05	0.16	0.07	0.14	-0.02	(-1.24)
Waste Collection	0.97	0.13	0.95	0.15	0.02*	(2.19)
Electricity	1.00	0.03	1.00	0.03	0.00	(0.51)
Water	0.95	0.16	0.95	0.13	0.00	(0.20)
Sewage	0.71	0.41	0.62	0.38	0.08*	(2.53)
Number of Students in High School	512.26	265.14	586.36	285.77	-74.10**	(-3.28)
Number of Female Students in High School	273.53	145.21	324.10	161.77	-50.57***	(-4.01)
GDP per Capita	18446.84	24039.44	13949.65	17123.75	4497.20**	(2.69)
City Population	68197.92	349673.41	148101.42	684066.28	-79903.50	(-1.75)
Mais Educação Federal Program	0.41	0.43	0.43	0.35	-0.02	(-0.69)
Number of Teachers in High School	0.00	0.00	0.03	0.43	-0.03	(-1.13)
Number of Teachers with High School Degree in High School	0.94	2.33	1.11	2.09	-0.18	(-0.98)
Number of Teachers with College Degree in High School	22.60	9.41	22.68	8.54	-0.08	(-0.11)
Proportion of High School Teachers with College Degree	0.96	0.10	0.95	0.09	0.01	(1.25)
Observations	334		273		607	

Note: School Census, ENEM, INEP, IBGE, MEC/FNDE, Administrative Data

APPENDIX B – Principal Component Analysis

As we approach in section 4.3.1, the mediators are high correlated with each other. This leads to multicolinearity problems, such as high variance and the dropping of some of the mediators in the estimation. Table 28 display the correlation matrix between mediators. For instance, mediator 1 and mediator 2 have a correlation of 99%. Mediator 3 and mediator 5 have a correlation of 99%. The information that these variables carry are highly redundant. It is likely that these variables are not really measuring seven different constructs.

Table 28 – Mediators correlation matrix

	Mediator 1	Mediator 2	Mediator 3	Mediator 4	Mediator 5	Mediator 6	Mediator 7	Mediator 8
Mediator 1	1							
Mediator 2	0.9933	1						
Mediator 3	0.8898	0.8978	1					
Mediator 4	0.6603	0.6558	0.7463	1				
Mediator 5	0.8969	0.8908	0.9926	0.7519	1			
Mediator 6	0.4166	0.4108	0.4791	0.3547	0.4858	1		
Mediator 7	0.5200	0.5234	0.4636	0.3353	0.4599	0.9484	1	
Mediator 8	-0.1484	-0.1546	-0.0607	-0.2172	-0.0537	-0.7743	-0.8393	1

Note: The mediators are dummies variables that correspond to the questions regarding the program's implementation. This table display the correlations between mediators.

To address this problem, we use the principal components analysis. This approach consists in reducing the variables into a smaller number of principal components that will account for most of the variance in the observed variables. We want to reduce these redundant variables into components that, in a lesser number, contain most of the variance of the variables.

This technique consists in (a) decomposing the correlation matrix of the mediators into eigenvectors. These vectors will be our components ; (b) selecting the eigenvectors that accounts for most of the variance of the original variables (in our case, mediators); (c) interpreting the components analyzing their correlation with the mediators.

The spectral theorem guarantees that the decomposition of the correlation matrix results in eigenvectors that are orthogonal between each other. With this, we have that the components are variables that explain most of the variance of the original mediators and are not correlated between them.

In table 29, we have the components resulting from the eigen decomposition, their respective eigenvalues, their proportion of the total variance of the mediators, and the cumulative variance. The components were ordered from the ones which account the most variance to the ones that account less. That is, Comp1 accounts for 0.5865 of the variance, Comp 2 accounts for 0.3349, and so on. We select only the first two components, Comp1

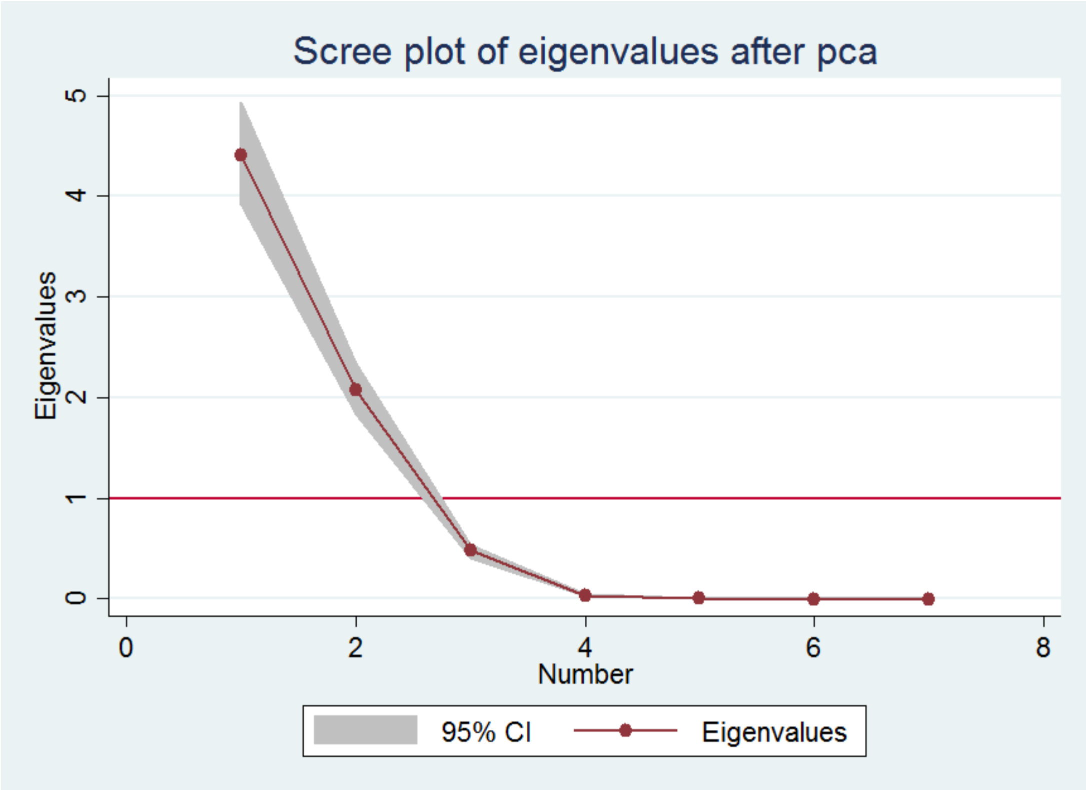
Table 29 – Eingenvalues

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	410.544	176.138	0.5865	0.5865
Comp2	234.407	186.485	0.3349	0.9214
Comp3	.479216	.438457	0.0685	0.9898
Comp4	.0407585	.0151614	0.0058	0.9956
Comp5	.0255971	.0206758	0.0037	0.9993
Comp6	.00492121	.00492121	0.0007	1
Comp7	0	.	0.0000	1

and Comp2. The two components have eingenvalues bigger than 1, and their cumulative variance accounts for 92.14% of the total variance. Figure 13 is the scree plot.

Having chosen the components, we proceed to their interpretation. Looking at the table 30, we have the correlation between mediators and Components. Note that Comp1 is highly correlated with the mediators 1 to 5. All of the are answers of questions related to political aspects of the implementation. For instance, mediator 1 is related to the engagement of the state governor in the implementation of the program. mediator 5 is

Figure 13 – Scree plot PCA



related to the implementation of the legal framework, that is, to the changes in law that are needed. Comp1, this way, will be addressed as political component. Comp2, for instance, is highly correlated to the mediators 6 and 7, that is, mediators related to the selection and removal process of teachers and principals. Therefore, we are going to address Comp2 as selection component. Since, we selected the two components altogether, and mediator 8 is not exactly related to these, we choose to use these components in our regressions, plus the mediator 8.

Table 30 – Components Mediator Correlation Matrix

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Unexplained
Mediator 1	0.4152	0.2750	-0.4721	0.3415	-0.3644	0.0228	0
Mediator 2	0.4161	0.2760	-0.4696	-0.1100	0.3964	-0.2757	0
Mediator 3	0.4084	0.2900	0.4762	-0.3445	0.4216	-0.0705	0
Mediator 5	0.4082	0.2896	0.4853	0.1665	-0.4442	0.2720	0
Mediator 6	0.3028	-0.4968	0.2584	0.5288	0.0962	-0.5530	0
Mediator 7	0.3393	-0.4680	-0.1269	0.1143	0.3499	0.7171	0
Mediator 8	-0.3380	0.4630	0.1075	0.6574	0.4508	0.1563	0

Note: Mediator 1 - "Was state governor engaged with the implementation of the program?", Mediator 2 - "Was the secretary of education engaged with the implementation of the program?", Mediator 3 - "Was there a dedicated team in Secretary of education for the implementation of the program?", Mediator 4 - "Did changes in the legal framework take place?", Mediator 5 - "Did all necessary changes in the legal framework take place, by the prescribed deadline?", Mediator 6 - "Did selection and removal of principals process take place?", Mediator 7 - "Did selection and removal of teachers process take place?", Mediator 8 - "Did selection and removal of teachers process take place?"