WORKING PAPER

#668 | 2024

Money for nothing and stigma for free? The effect of positive discriminatory policies on education gaps

JOSÉ MESQUITA GABRIEL, LUÍS CATELA NUNES

NOVA SCHOOL OF BUSINESS & ECONOMICS

Money for nothing and stigma for free? The effect of positive discriminatory policies on education gaps[1]

José Mesquita Gabriel[2]
Nova School of Business and Economics

Luis Catela Nunes[3]
Nova School of Business and Economics

Abstract

This paper analyzes a compensatory education program in Portugal whose aim is to provide equal educational opportunities for children from lower socioeconomic status (SES) families by ensuring additional school resources for deprived areas. Using two administrative databases covering virtually all public schools in Portugal, we present a comprehensive evaluation of such a program addressing an important effect that has been overlooked so far in the literature: a negative stigma effect, wherein students from higher SES families become less likely to enroll in *treated* schools. Our staggered difference-in-differences estimates show that: 1) the *Student-to-Teacher* ratio decreased significantly in *treated* schools as a result of their entry into the program, corresponding to an average drop of 9% relative to the pre-treatment average value; 2) the proportion of *students whose mothers concluded upper secondary school* entering *treated* schools declined substantially, a decrease of around 14% of the average pre-treatment value; and 3) no effects were observed in blind-marked national exam scores in the 4th-, 6th-, and 9th-grade for students coming from lower SES families, while some positive effects were found for non-blind sources of evaluation, particularly in schools where the change in additional resources was more pronounced. Our results emphasize the need to account for unanticipated risks of further aggravating segregation across schools when implementing publicly announced programs, in particular when they lead to discontinuities in terms of school eligibility.

Keywords: Education; Compensatory education programs; Inequalities; Portugal.

JELs: C21, H52, I24, I28

---

[1] This project was funded by Fundação para a Ciência e Tecnologia (PTDC/EGE-ECO/4764/2021, UIDB/00124/2020, UIDP/00124/2020 and Social Sciences DataLab - PINFRA/22209/2016), POR Lisboa and POR Norte (Social Sciences DataLab, PINFRA/22209/2016). The authors thank DGEEC for providing access to anonymized education microdata for research purposes. Any errors are our own.

[2] Nova School of Business and Economics, Universidade Nova de Lisboa, Campus de Carcavelos, 2775-405 Carcavelos, Portugal. Email: jose.gabriel@novasbe.pt.

[3] Nova School of Business and Economics, Universidade Nova de Lisboa, Campus de Carcavelos, 2775-405 Carcavelos, Portugal. Email: lcnunes@novasbe.pt.

I. Introduction

The relationship between school spending and student attainment has long been a highly debated topic for researchers and policymakers worldwide. Hanushek (2003) presented a first revision of the US-based literature and showed that the majority of studies to date had not uncovered any significant positive impact of an increase in school resources on student academic performance. However, with the gradual increase in the availability of disaggregated administrative data, recent studies with sound identification strategies in multiple quasi-natural experiments have uncovered a positive causal relationship between school inputs and student outcomes (see Jackson 2020 for a detailed survey).

Compensatory education programs are incorporated into this branch of literature and focus on students from a lower socioeconomic status. These programs provide schools in less advantageous socioeconomic areas with more funds per student to counter existing academic inequalities and ensure that less privileged children have the same educational opportunities as their peers.

Despite its importance across several developed countries, acting as the main policy to reduce SES achievement gaps, previous empirical evaluations of such programs are still rather scarce and have produced mixed results so far (for a comprehensive review, see Franck and Nicaise, 2022).[4] On the one hand, some programs, particularly those targeting students at earlier stages of education, have been associated with positive outcomes: in the US, Johnson and Jackson (2019) found that increases in pre-school spending through the *Head Start* program increased individual educational attainment and lifetime earnings, while Carneiro and Ginja (2014) found that participation in the program reduced the incidence of behavioral and health problems; in the UK, Machin et al. (2010) examined the *Excellence in Cities* program and uncovered positive causal effects in Mathematics scores and school attendance among high-ability students in disadvantaged schools.[5] On the other hand, studies have also found that compensatory education programs may be ineffective in improving test scores: in the US, Van der Klaauw (2008) and Matsudaira et al. (2012) examined the *Title I* program, a federal funding policy aimed at ensuring high-quality education for disadvantaged pupils, and both drew parallel conclusions that the program was ineffective in improving test scores; in France, Bénabou et al. (2009) evaluated the French program *Zones d'Education*

[4] Examples of such place-based educational policies include the *Title I* program (US), the *Education Action Zones* (UK), or the *Zones d'Education Prioritarie* (France and Belgium).

[5] In the Netherlands, De Haan (2017) also uncovers a positive effect of the *Learning Support* program on students' passing rates; unlike most compensatory education programs, the focus was on students lagging behind, regardless of their socioeconomic status.

Prioritarie and showed that the program was successful in providing some additional teaching resources to schools with underprivileged pupils, even though no effects on student achievement were uncovered.[6]

As such, we view our research focused on the Portuguese compensatory education program – the *TEIP* program – as novel and an important contribution to the literature for two key reasons. First, by taking advantage of two administrative databases containing information on virtually every student in every public school in Portugal, we uncover a largely overlooked indirect impact of the program: a negative stigma effect towards participating schools, characterized by a significant decrease in the proportion of students from higher SES families enrolling in these schools. Second, we present, for the first time, estimates of the effect a compensatory education program has on student educational achievement throughout different educational stages when accounting for this significant compositional effect.

Previous literature taking advantage of particular thresholds defining eligibility for programs and comparing the evolution of schools close to a certain cut-off point might have produced biased estimates of the impact that such programs have on student achievement, particularly on those the policy was originally designed for – low SES students, who are less likely to leave participating schools.

Our identification strategy exploits the fact that the program, which accounts for around 15% of the total Portuguese student population in basic education, had three distinct phases – the implementation of the program, in 2006 (*TEIP1*); the first expansion in 2009 (*TEIP2/3*); and the second one, in 2013 (*TEIP4*).[7] As such, we employ a *staggered difference-in-differences* methodology using the different implementation phases to evaluate the program.

Our results show, first and foremost, that additional resources were, in fact, provided to *treated* schools – we estimate a statistically significant drop in both the S*tudent-to-Teacher ratio* and the *Student-to-Staff ratio* of around 8% of the average pre-treatment value for treated units. Importantly, this effect has not always been found in previous literature.

We also find evidence of a significant negative sorting effect, by which more informed parents opt to enroll their children elsewhere once it becomes public that a particular school has a large percentage of students from an unfavorable socioeconomic

[6] In Portugal, the vast majority of studies have been qualitative oriented and based on a small number of schools (see Fritsch and Leite, 2020, for an overview).

[7] Two important notes, which are mentioned in more detail later – 1) the program had a short-lived pilot project in the 1990s; and 2) we opted to tag each stage of the program according to the specific expansion phase. Officially, *TEIP1* refers to the pilot project period; *TEIP2* includes the period of adherence of our *TEIP1* and *TEIP2/3* schools; and *TEIP3* corresponds to the period where our *TEIP4* also entered the program.

background. In fact, we estimate that the proportion of *students whose mothers concluded upper secondary school* entering *treated* schools significantly declined after treatment by around 3.7 p.p. (14% of the average pre-treatment value for treated units). We show that this effect is also observed for different higher education attainment levels, both for the students' mothers and fathers.

Given this sorting effect, we examine the impact on student achievement in two main sub-samples: *students whose mothers have not concluded upper secondary school* and *students in the highest bracket for receiving social support from the government – Type A*. By doing so, we 1) account for the negative stigma effect, as these are students whose financial constraints significantly hamper their ability to change schools, and 2) focus our analysis on the students that the program was intended for.

Irrespective of the sub-sample we focus our attention on, our results are highly consistent across the board – we do not find any statistically significant effect of the program on average student scores in blind-marked standardized national exams in the 4^{th}-, 6^{th}-, and 9^{th}-grade. Importantly, and as a result of how similar our *treated* and control groups are, we show that both groups revealed remarkably parallel development before treatment, a result that supports the interpretation of our estimates as a causal effect of the program.

Furthermore, we also examine internal sources of evaluation carried out by schoolteachers. Here, our findings are more heterogeneous – while retention rates in basic education remained unaffected, we show that retention rates dropped in lower secondary education, and this effect is driven by schools where the change in additional resources was more pronounced.

Our study sheds light on the importance of evaluating public policies to ensure public resources are used most efficiently. In particular, our study emphasizes the unanticipated risks that may arise from publicly announced positive discriminatory programs in further aggravating segregation across schools.

The remainder of this paper is organized as follows. Section 2 presents the institutional background. Section 3 presents the data and methodology employed, while Section 4 presents the results. Section 5 is the conclusion.

II. Institutional Background

II.1. The Portuguese educational system

The Portuguese educational system is divided into four main levels: pre-primary, basic, secondary, and tertiary education.[8] Pre-primary is not mandatory. Children must enroll in the 1st grade of basic education in the year they turn 6.[9] Education is compulsory until the age of 18 when students usually complete the 12th grade. The majority of this time is spent in basic education, which is divided into three different cycles: the first cycle corresponds to grades 1-4 (usually, from ages 6 to 9); the second cycle corresponds to grades 5-6; and the third cycle comprises grades 7-9.

In order to finish basic education and proceed to secondary school, students must undertake two high-stakes standardized national exams at the end of the 9th grade, in Portuguese and Mathematics. Until the 2014/2015 school year, students were also required to take two national exams in the 4th and 6th grades, also in these two subjects.

After completing basic education, students decide whether to follow the vocational or academic track at secondary school, the latter intended for those who wish to pursue tertiary education, while the former is usually chosen by students aiming to enter the labor market earlier.[10]

Across all levels of education, students have access to both publicly and privately funded schools, even though the majority of students attend the former – during the academic years analyzed (2006/2007 – 2017/2018), according to the official counts, the percentage of students enrolled in public schools in basic and secondary education was approximately 85% and 80%, respectively.

II.2. The *TEIP* program

After a short-lived pilot project in the 1990s, the *TEIP* program was re-established in the 2006/2007 school year with the same 35 school clusters divided across the *Lisbon* and *North* regions (*TEIP1*). Since then, the program has been progressively growing.

The program saw its first expansion during the 2009/2010 school year: at the start of the school year, 24 school clusters adhered to the program, and later in the same academic year, 45 more became part of the program (*TEIP2*). Consequently, at the end of the

[8] In Portuguese: "Educação pré-escolar, ensino básico, ensino secundário e ensino superior".
[9] Even though pre-primary is not mandatory, during the period analyzed (2006/2007 – 2017/2018) around 90% of children of eligible age were enrolled.
[10] For a graphical representation of the Portuguese educational system, see figure A1.

2009/2010 school year, the number of school clusters had practically tripled to 104, most of which were located in the *Lisbon* and *North* region, albeit with some representation all over the Portuguese mainland. The fourth expansion phase occurred in the 2012/2013 school year, with the inclusion of 33 school clusters in the program (*TEIP3*). It is important to note that no school cluster has left the program; as such, and by the 2017/2018 school year, a total of 137 school clusters were part of the program, accounting for around 15% of the total student population in basic education.[11]

The successive enlargement of the program is a consequence of the political strategy delineated by different Portuguese governments over the years, as the inclusion of school clusters in the *TEIP* program depends on an invitation made by the Ministry of Education. In fact, the Ministry defines which schools are invited to the program by identifying schools where a large percentage of students are recipients of social support from the government, are from an immigrant background, and where a large percentage of students' mothers have not concluded secondary school.[12] Once the invitation is made, the school cluster decides to accept or reject it, knowing that adherence to the program implies increased funding, both from the Ministry of Education and from the European Union (through the EU cohesion policy), and autonomy concerning where to invest the additional funds – whether on teaching or non-teaching staff.[13]

III. Data and Methodology

III.1. Data

To perform our study, we take advantage of a comprehensive administrative database (MISI) made available by the Directorate General of Education and Science Statistics

[11] At the beginning of the 2021/2022 academic year, the program saw its 5th expansion phase, with the addition of 10 new school clusters, making a total of 146 (two *TEIP* school clusters were merged into one). Furthermore, starting in the 2024/2025 school year, a new wave of expansion is taking place, with the adherence of up to 24 school clusters. However, due to data limitations, this is outside the scope of our study.

[12] These are the established priorities but no known threshold or transparent formula defined invitation to the program. Only in the upcoming expansion phase, occurring during the 2024/2025 school year, have transparent weights been given to each variable: each school cluster is ranked according to the weighted percentile of a) the percentage of students receiving social support (weighted 50%), b) the percentage of students whose mothers have not concluded secondary school (25%), and c) the percentage of students from an immigrant background (25%) (for a more detailed description of the upcoming expansion phase, see the following link, in Portuguese: https://dge.mec.pt/sites/default/files/Noticias_documentos/aviso_dge_teip4_janeiro2024.pdf).

[13] More than 80% of the funds are incorporated into the Human Resources category and include the hiring of teachers, psychologists and mediators, among others. The remainder goes into the Acquisition of Goods and Services category, encompassing the hiring of specialized technicians, cultural facilitators and expenses on books and meals not covered by other programs.

(DGEEC), which comprises extensive information on every student and staff member in every public school in mainland Portugal between the 2006/2007 and 2017/2018 school years. With this dataset, it is possible to keep track of a student throughout their entire academic career until the end of secondary school. We also rely on a dataset provided by the Portuguese Ministry of Education with information regarding scores in national exams in the 4th, 6th, and 9th grades (ENEB).[14]

With the complete merged datasets, we have information on student academic performance, both from school evaluation (internal grades) and from blind-graded standardized national exams. Furthermore, we have access to several socio-demographic variables, such as age, gender, nationality and place of birth, whether the student has access to a computer or the internet at home, and whether the student receives social support from the government (and if so, in which bracket).[15] We also have information about students' parents, such as educational level, nationality and employment status.[16]

Information on students enrolled in private schools is also available but more limited regarding socioeconomic characteristics. However, this information is not crucial in our study for two main reasons. The first is that more than 85% of students enrolled in basic education in Portugal attended public schools during the period analyzed. The second is that, on average, students from private schools come from more favorable socioeconomic backgrounds and thus should not be used as a counterfactual for students in *TEIP* schools.

Given our research question, we can aggregate all outcomes of interest used during our study as follows. Regarding whether the program brought additional resources to schools, we investigated the evolution of the *Student-to-Teacher* and *Student-to-Staff* ratios, both measured at school level. To analyze whether the composition of students in schools changed, we examined the characteristics of students entering *treated* schools in the 5th and 7th grades, namely their parents' education and their exam scores from the previous year. Finally, we employ several measures of student achievement from the 1st to the 9th grades. We analyze the evolution of blind-graded national standardized exams in the 4th, 6th, and 9th grades. We examine both average scores and passing rates. And we assess three indicators to

[14] Our final dataset has observations from around 10.2 million students and 2.1 million staff members over 12 academic years, where each observation corresponds to one student/staff member in one given year.
[15] In Portugal, low-income students are entitled to school subsidies to support their expenses with meals at the school, school supplies and textbooks. These expenses can be fully (Bracket A) or partially subsidized (Bracket B), depending on the parents' income levels.
[16] The anonymized datasets used in this study are available from DGEEC. Descriptive statistics for all the relevant variables are presented in Table A1.

examine internal sources of evaluation, namely whether the student passed the *7th and 8th grades without retentions*, whether they passed the *5th and 6th grades without retentions* and whether they showed *successful progress in lower secondary education*, which implies no retentions during the 7th and 8th grade, as well as passing both 9th grade national exams.

III.2. Methodology

The central aim of this study is to understand what the impact of the Portuguese compensatory program has been, specifically 1) whether it has, in fact, been able to provide additional resources to schools; 2) whether it has led to a change in the characteristics of students entering these schools due to stigmatization of targeted areas; and 3) whether it has fostered academic achievement, particularly for students from a lower socioeconomic status.

Recently established literature has pointed out the problems with employing the traditional two-way fixed-effects (TWFE) model in treatments under a staggered implementation, as is the case of the *TEIP* program. The work done by researchers such as Goodman-Bacon (2021), Callaway et al. (2024), and de Chaisemartin and D'Haultfœuille (2020) has shown how "forbidden comparison" units are used in the traditional TWFE estimation procedure, leading to biased estimates under dynamic treatment effects.

For this reason, we employ as our preferred method the one proposed by Callaway and Sant'Anna (2021), which consists of estimating cohort-time-specific treatment effects without problematic control units.[17] The key identifying assumption is that had schools not adhered to the program, the expected evolution of student educational outcomes would have been the same as that observed in the control group, conditional on covariates (conditional parallel trends assumption). This assumption is discussed in detail and supported with different arguments in each subsection of the results.[18]

Another important identifying assumption is the limited treatment anticipation, i.e., that schools did not change their behavior before entering the program, anticipating that this would occur later on. In the present study, this is not a cause for concern given how admission to the program is defined and announced only after the beginning of the school year that precedes adherence.

[17] Nevertheless, we also show that our results are robust to other methods, such as the one proposed by Sun and Abraham (2021) or the canonical TWFE.
[18] Furthermore, tests were also carried out on whether the trends in the treated and control groups were similar before the start of the TEIP program. The p-values of pre-trend tests, based on chi-squared statistic, of the null hypothesis that all pre-treatment ATTs are equal to zero, are shown in the estimation results' tables throughout the study.

We define treatment groups according to the year g in which they entered the program. The average treatment effect for a particular group g in a particular year t is denoted by the difference in a certain outcome of interest, in expectation terms, between year t and baseline year $g-1$ (the year before treatment) across units in *treated* group g and comparison group $\mathcal{G}_{comp}$, as presented below:

$$ATT(g,t) = \mathbb{E}[Y_{it} - Y_{i,g-1} \mid \mathcal{G}_i = g] - \mathbb{E}[Y_{it} - Y_{i,g-1} \mid \mathcal{G}_i = \mathcal{G}_{comp}]$$

where $\mathcal{G}_i$ denotes the group which a unit i belongs to and the comparison group, $\mathcal{G}_{comp}$, includes both never-treated and yet-to-be-treated units.[19]

In order to aggregate the different ATTs and have estimates with a similar interpretation as event-study regressions, we let e denote time since treatment was adopted in a certain group g, i.e., $e = t - g$, and consider the following aggregation:

$$\theta_{es}(e) = \sum_{g \in \mathcal{G}_{treated}} ATT(g, g+e) * \mathbf{1}\{g + e \leq \mathcal{T}\} * P(\mathcal{G}_i = g \mid \mathcal{G}_i + e \leq \mathcal{T})$$

where $\theta_{es}(e)$ is the average effect of participating in the program e periods after implementation and $\mathcal{T}$ represents the last year analyzed (in our case, the 2018/2019 school year). This way, we can distinguish between the program's impact in the short- and medium-term.

Furthermore, to present a more aggregate causal parameter, for each specification we present the average treatment effect on the treated with an analogous interpretation as in the canonical DiD setup, as follows:

$$\theta_w^0 = \frac{1}{k} \sum_{g \in \mathcal{G}_{treated}} \sum_{t=2007}^{\mathcal{T}} ATT(g,t) * \mathbf{1}\{t \geq g\} * P(\mathcal{G}_i = g \mid \mathcal{G}_i \leq \mathcal{T})$$

where $k = \sum_{g \in \mathcal{G}_{treated}} \sum_{t=2007}^{\mathcal{T}} \mathbf{1}\{t \geq g\} * P(\mathcal{G}_i = g \mid \mathcal{G}_i \leq \mathcal{T})$, to ensure that the weights of all $ATT(g,t)$ add up to one. As such, θ_w^0 is simply a weighted average, according to group size, of each $ATT(g,t)$. This implies that larger schools and schools that adhered to the program earlier will be given more weight vis-à-vis the remaining schools. Importantly, this way we assign the same weight for each treated observation, that is, for each student in a certain year t enrolled in a school that is part of the program.

[19] We make use of the STATA commands csdid (Rios-Avila et al., 2023) and eventstudyinteract (Sun, 2022).

To account for possible correlations of the outcome variable between students within each school cluster and as treatment assignment is defined at such a level, the statistical inference uses cluster-robust standard errors at school cluster level (see Abadie et al. 2023).

IV. Results

IV.1. Were additional resources provided to schools?

The central aim of compensatory education programs is to provide equal educational opportunities for children from a lower socioeconomic status by providing targeted areas with additional school resources. As such, we start by examining whether adherence to the program effectively translated into extra resources for *treated* schools.

We start by analyzing the *Student-to-Teacher* ratio at the school level. In Figure 1, we show that this ratio significantly decreases after treatment, a clear indication that the *TEIP* program has effectively provided additional resources per student.[20] Our estimates uncover an average treatment effect on the treated of -0.79. For a clearer understanding of the magnitude, this corresponds to 8.7% of the pre-treatment average of the *Student-to-Teacher* ratio for *treated* schools.[21]

Along with additional resources, schools are also granted a significant degree of autonomy. In fact, the choice of where to invest the extra funds is made at school cluster level and schools can choose whether to hire more teaching or non-teaching staff. We examine how schools have made such decisions by further analyzing the *Student-to-Staff* ratio at school level.[22] Our results uncover a significant decrease in this ratio following admission into the program – we estimate an average effect of -0.56, which corresponds to a decrease of around 8.2% of the pre-treatment average for treated schools. Given the similarity between both estimates, schools seem to have invested proportionately in both teaching and non-teaching staff, yet slightly more in the former.

In both analyses, it is reassuring to note the absence of any pre-treatment trend, providing support to a causal interpretation of the estimation results when employing a

[20] More information about the point estimates are presented in Table A3.
[21] Figure A4 examines the changes in *Adjusted Student-to-Teacher* ratio, where teachers not assigned to a particular school, but rather to a school cluster, are accounted for. A similar effect is estimated.
[22] Staff includes both teaching and non-teaching personnel.

difference-in-differences approach. As such, the evidence is consistent with schools having additional human resources as a direct result of adherence to the program given that the evolution of both ratios was very similar between treated and non-treated schools before the program's implementation, and it seems rather unplausible that it is the consequence of another confounding effect.

For a clearer understanding of the main impetus behind the decrease in the ratios, we disentangle the effect coming from the number of students enrolled and teachers employed at *treated* schools. In Figures A5 and A6, we show that the implementation of the program has a significant negative impact on the number of students (around 55 fewer students per school, or 10.4% of the pre-treatment average), while there is no significant impact on the number of teachers. Importantly, as schools are granted resources according to the number of students enrolled, the fact that the number of teachers remained constant as the number of students decreased represents the estimated causal impact in the number of resources per student as a direct effect of the program.

Regarding heterogeneity across different *TEIP* expansion phases, Figure A7 and Table A3 illustrate that the decline in the *Student-to-Teacher* ratio is more pronounced for *TEIP2/3* schools, where we estimate a statistically significant effect of -0.98 (or 10% of the pre-treatment average), compared to *TEIP4* schools, where we identify an impact of -0.25 (or 3%).

Our results point to a clear effect of the program in being able to provide extra resources per student in targeted schools, with the decrease in the *Student-to-Teacher* and *Student-to-Staff* ratio corresponding to almost 10% of the pre-treatment average for treated units, in line with previous literature on compensatory education programs.

As a comparison, Bénabou et al. (2009) found that the French *ZEP* program involved an extra 10% in resources per student, while Machin et al. (2010) estimate that the average additional resources brought by the *Excellence in Cities* program corresponds to £120 per student per year (almost 5 percent of the overall per pupil expenditure). For the *Title I* program in the US, Van der Klaauw (2008) does not find statistically significant changes in per-student expenditures, with a possible explanation being that some states shifted funding from treated to untreated schools once the program was implemented (Gordon, 2004).

Furthermore, we show that schools have employed the additional resources to hire both teaching and non-teaching personnel at a comparable rate and that the group of schools in the *TEIP2/3* expansion was associated with a larger decrease in the *Student-to-Teacher* ratio vis-à-vis *TEIP4* schools.

IV.2. Has the composition of students changed?

Place-based compensatory programs aim to improve student achievement through positive discriminatory measures. However, the public distinction of which schools are being targeted might raise awareness regarding the proportion of students from a less advantageous socioeconomic background, making it a more salient topic (Bordalo et al., 2022). In turn, a negative stigma might lead some of the parents, possibly the more informed ones, to enroll their children in non-targeted schools, further exacerbating segregation and hindering student achievement.[23]

To quantify this possible side effect of the policy, we examine whether *treated* schools were associated with an increase in the proportion of students coming from a lower socioeconomic background and with worse previous academic performances after the program implementation *vis-à-vis* non-treated schools. We focus on students entering schools in the 5^{th} and 7^{th} grades, as these are the two moments when students usually change schools during basic education in Portugal.

Our results show that the fraction of students from high SES families entering *treated* schools seems to have significantly decreased. In fact, in Figure 2, we find that the proportion of *students whose mothers concluded upper secondary school* significantly declines by around 3.7 p.p., corresponding to a 14% decrease compared with the pre-treatment mean for treated units.[24]

Notably, we observe that the effect appears to increase over time, consistent with the models of dynamic segregation of Schelling (1969, 1971) and the concept of tipping points discussed in Caetano and Maheshri (2017). Actually, a self-fulfilling prophecy may develop whereby even parents who initially do not alter their perception of the proportion of low SES students, eventually notice a *de facto* change in the school's socioeconomic composition and decide to stop enrolling their children in a treated school, even though the initial effect may be the result of some parents overreacting to the news (Bordalo et al., 2020).

[23] The increase in school segregation as a result of an increase in immigrant inflows (Betts and Fairlie, 2003) or the number of minority schoolchildren (Fairlie and Resch, 2002), sometimes referred to as "white flight", has been well documented.

[24] Figure A8 uncovers a similar negative effect in the proportion of *students whose mothers concluded a bachelor's degree*. In Figures A9 and A10, we find no effect on the proportion of *students in the highest bracket for receiving social support from the government – Type A* and those from an immigrant background entering *TEIP* schools, respectively.

Furthermore, it is reassuring to observe common pre-trends in Figure 2, i.e., schools in the treatment and control group were developing identically in the pre-treatment period and, as such, it is improbable that this would change later in the absence of treatment. Nevertheless, given that the schools targeted were in areas with lower levels of parental educational attainment, it can simply be the case that municipalities where *treated* schools were located had a distinct evolution in terms of educational attainment, contemporaneous with but not the consequence of the program implementation. For this reason, we ran a robustness test that only compares students in the same municipality by adding year × municipality fixed-effects, and we found similar results in Figure A11.

Another central aim of this section is to understand whether the compositional effect is specific to socioeconomic background or is also present in terms of student abilities, as measured by their previous exam scores. For this reason, we analyze the evolution of blind-graded exam scores in the academic year before entering the 5th and 7th grades.

Our results, presented in Figure 3, show that there was no effect in terms of the average scores of students entering treated *vis-à-vis* non-treated schools in the 5th grade when measured by the *4th grade Mathematics exam score*. However, when we focus our analysis on the *6th grade Mathematics exam score* obtained by students entering the 7th grade the following year, we see a statistically significant decrease in the score, although one of lower magnitude – we estimate an average effect of -0.09 (on a scale of 1-5), which translates to around 2 points on a scale of 0-100.[25]

Overall, the results presented in this section shed light on an important negative stigma effect that has been largely unaccounted for in the scarce literature evaluating compensatory education programs so far – parents from a higher socio-economic background, conceivably more informed about the program and about which schools have been selected for the program, become less likely to enroll their children in targeted schools, further exacerbating segregation in the more deprived areas.[26]

[25] Figures A12 and A13 show an identical scenario when analyzing Portuguese exam scores in the 4th and 6th grade, respectively.

[26] An exception is the study by Davezies and Garrouste (2020), which also uncovers a negative stigma effect in the French *RAR* program – although only examining two student cohorts.

IV.3. Were there changes in academic achievement?

In the last part of this section, we focus on our central issue, which is to understand whether the program boosted academic achievement among students from a lower socioeconomic background and, thus, helped bridge achievement gaps.

As a consequence of the results uncovered so far regarding the negative stigma effect, we have decided to primarily focus our analysis on two sub-samples: *students whose mothers have not concluded upper secondary school* and *students receiving the highest level of social support from the government – Type A*. By doing so, we place a greater emphasis on the subset of students for whom the program was developed.[27] Furthermore, this allows us to account for the aforementioned compositional effect, as these are students whose financial constraints significantly hamper their ability to switch schools. As such, we compare the evolution of similar students (both regarding socioeconomic and academic characteristics, as shown in Table A2) in treated and not-treated (or yet-to-be-treated) schools.

We emphasize that, unlike previous literature, we estimate the average treatment effect on targeted students, that is, low SES students; thus, our methodology differs from the ones in studies using certain thresholds or cut-off points defining eligibility to a program, and consequently comparing the average educational attainment of all students in schools that were just accepted to the program vis-à-vis in schools that were just rejected.

As mentioned before, we examine the evolution of both blind-graded exam scores and internal scores given by schoolteachers. We start our analysis with the former in Table 1. We show that irrespective of the sub-sample we focus on, for all exams taken in the 4th, 6th, and 9th grades, we uncover an effect on scores that is not statistically significantly different from zero.[28]

Once again, it is encouraging to note that for all outcomes considered in Table 1, we never reject the null hypothesis that all pre-treatment ATTs, by group and year, are equal to zero. This provides further support to interpret the results presented in this table as a direct effect of the program, as 1) the groups were similar before treatment and 2) they showed

[27] We emphasize that, unlike previous literature, we estimate the average treatment effect on *targeted* students, that is, low SES students; thus, our methodology differs from the ones in studies using certain thresholds or cut-off points defining eligibility to a program, and consequently comparing the average educational attainment of all students in schools that were just accepted to the program vis-à-vis in schools that were just rejected.

[28] At the usual 5% level. At the 10% level, we uncover a small positive effect on the Mathematics 4th grade exam scores for the *students whose mothers have not concluded upper secondary school* and a negative effect on the Portuguese 9th grade exam scores for *students in the highest bracket for receiving social support from the government – Type A*.

statistically indistinguishable evolution in pre-treatment years.[29] Furthermore, Tables A4 and A5 show that our results are consistent when using other methods, namely the method proposed by Sun and Abraham (2021) and the traditional TWFE, respectively.

We complemented our examination of external sources of evaluation by also employing the percentage of passing scores obtained in each exam as an outcome variable. The results shown in Table 2 are similar across the board – whether we look at the aggregate result of the three different moments of evaluation in Panel A, at the results by *TEIP* expansion phase in Panel B, or at the results by educational stage in Panel C, we always uncover an effect statistically indistinguishable from zero, in line with our findings from Table 1.

Finally, we analyze internal sources of evaluation. In Panel A of Table 3, we show that the probability of a student achieving *successful progress in lower secondary education*, defined as passing both the 7th and 8th grades without retentions and having a passing score in both 9th grade national exams, was unaffected by the program. This is the case irrespective of the sub-sample employed. Furthermore, in Panel C in the same table, we show that the probability of passing the *5th and 6th grades without retentions* was not impacted by adherence to the *TEIP* program either.

However, we do uncover a statistically significant positive effect when examining the proportion of students who pass the *7th and 8th grades without retentions.* In panel B of Table 3, we show that the probability of passing the *7th and 8th grades without retentions* significantly increases by around 3 p.p. in both sub-samples employed (or around 5% of the pre-treatment average for treated units).[30] Moreover, we show that this effect is entirely driven by *TEIP2/3* schools, the expansion-phase group associated with a more pronounced increase in additional resources.

As such, we show that the program's overall impact on student academic achievements was null when analyzing national exam scores, both at primary and lower secondary levels (basic education). Furthermore, we only uncover a positive effect in internal sources of evaluation in lower secondary education, prompted by the expansion-phase group associated with a larger increase in additional resources (both teaching and non-teaching staff) brought to schools.

[29] For an event-study graph, see Figures A14-A19.
[30] See Figures A19-A22 for the presentation of the event-study estimates.

Overall, our estimates showing a null impact of the program on student academic achievements seem to be the consequence of two conflicting effects. On the one hand, the provision of more resources per student. In fact, studies using the timing of court-mandate reforms as a source of exogenous variation in school spending have shown significant positive effects on student achievement: in the US, Jackson et al. (2016) showed that a 10 percent increase in per-pupil spending, a similar order of magnitude as the *TEIP* program, led to an increase in completed years of education and higher wages (more 0.27 years and 7.25 percent, respectively), particularly among low-SES students; Candelaria and Shores (2019) found that seven years after court-mandated reforms, a per-pupil spending increase of around 12 percent translated into a significant rise in graduation rates, of between 6 and 12 percentage points; and Hyman (2017) showed that Michigan's *Proposal A* reform led to an 8.4 percent decrease in class size relative to the mean, which in turn increased the likelihood of treated students enrolling in college by around 3 percentage points.[31] These studies, in line with a meta-analysis performed by Jackson et al. (2024) that showed that reforms in public school spending tend to be more effective for economically disadvantaged populations, are a good indicator of an upper-bound on the beneficial impact of extra resources made available by compensatory education programs on academic achievement.

However, the compositional effect whereby students enrolled in *treated* schools become associated, on average, with a higher percentage of peers from a low SES background – which has been found in the literature to have a substantial negative impact on student achievement (see Van Ewijk and Sleegers, 2010, for a meta-analysis) – might counteract the positive effects stemming from extra funding to disadvantaged areas.

[31] More generally, some studies using quasi-natural experiments have reported sizeable effects of reducing class size (e.g., Angrist and Lavy, 1999; Haegeland et al., 2012), while other recent studies have uncovered smaller impacts (e.g., Cho et al., 2012), particularly regarding extra funding used in computers and software (Leuven et al., 2007). For a meta-analysis, see Krueger (2003).

V. Conclusion

Our study contributes to the scarce literature on compensatory education programs by showing an important negative stigma effect that has been vastly unaccounted for so far – parents from a higher socioeconomic background become significantly less likely to enroll their children in targeted schools after implementation of the program, thus further exacerbating segregation in different areas.

This compositional effect is crucial when assessing the impact on educational achievement for those the program is designed for – students from lower SES families. Failing to account for this effect will lead to a downward-biased estimate of the program, as students leaving treated schools are associated with above-average inputs. As such, we consider this effect and carry out an analysis at the student level for the first time, focusing specifically on students from lower SES families rather than performing a school-level analysis.

Employing an administrative database to evaluate the Portuguese compensatory education program (*TEIP*), our findings show that additional resources were, in fact, provided to schools – the *Student-to-Teacher ratio* and the *Student-to-Staff ratio* significantly decreased after the implementation of the program (a drop of around 8% of the average pre-treatment value). Furthermore, the aforementioned negative stigma effect arose for adherents to the program – the proportion of *students whose mothers concluded at least upper secondary school* entering *TEIP* schools significantly declined after treatment, by around 3.7 p.p. (14% of the average pre-treatment value).

Our *staggered difference-in-differences* estimates show that after considering this compositional effect and focusing on the sub-sample of students from lower SES families, student academic achievement remained practically unchanged at basic education level. This is the case when analyzing blind-marked external sources of evaluation, namely the 4th-, 6^{th,}-, and 9th-grade standardized national exam scores. Regarding internal sources of evaluation given by schoolteachers, we found some heterogeneity – while some indicators show no changes, we show that retention rates fell in lower secondary education, and this effect is driven by schools where the change in additional resources was more pronounced. We note, however, that different dimensions of student achievement, namely effects on dropout rates, school attendance, and violence at school, are important in determining the success of the program but were not assessed owing to data limitations; future research exploring these dimensions is thus necessary to fully comprehend the impact of the program.

Our results have important policy implications – they highlight the costs of having a publicly announced program binarily targeting and labelling schools that have a large percentage of students from lower SES families. One way to counter the stigmatization of these areas while bringing in the additional resources needed to ensure equal educational opportunities would be either to adjust school funding proportionally according to the percentage of low SES students, thus in a *continuous* rather than *binary* manner or, in case there is a concern that funds could be misallocated to higher SES students enrolled in the school, to channel the additional resources to students directly, rather than to schools.

References

Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics*, 138(1): 1-35. https://doi.org/10.1093/qje/qjac038

Angrist, Joshua D., and Victor Lavy. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *The Quarterly Journal of Economics*, 114(2): 533–575. https://doi.org/10.1162/003355399556061

Bénabou, Roland, Francis Kramarz, and Corinne Prost. 2009. "The French zones d'éducation prioritaire: Much ado about nothing?" *Economics of Education Review*, 28(3): 345-356. https://doi.org/10.1016/j.econedurev.2008.04.005

Betts, Julian R., and Robert W. Fairlie. 2003. "Does immigration induce 'native flight' from public schools into private schools?" *Journal of Public Economics*, 87(5-6): 987-1012. https://doi.org/10.1016/S0047-2727(01)00164-5

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *American Economic Review*, 110 (9): 2748–82. https://doi.org/10.1257/aer.20181219

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2022. "Salience." *Annual Review of Economics*, 14(1), 521-544. https://doi.org/10.1146/annurev-economics-051520-011616

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "*Difference-in-differences with a continuous treatment.*" National Bureau of Economic Research. https://doi.org/10.3386/w32117

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of econometrics*, 225(2): 200-230. https://doi.org/10.1016/j.jeconom.2020.12.001

Caetano, Carolina, Brantly Callaway, Stroud Payne, and Hugo S. A. Rodrigues. 2022. "Difference in Differences with Time-Varying Covariates." *arXiv preprint arXiv:2202.02903*. https://conference.iza.org/conference_files/EVAL_2022/callaway_b32685.pdf

Caetano, Gregorio, and Vikram Maheshri. 2017. "School segregation and the identification of tipping behavior." *Journal of Public Economics*, 148: 115-135. https://doi.org/10.1016/j.jpubeco.2017.02.009

Candelaria, Christopher, and Kenneth A. Shores. 2019. "Court-ordered finance reforms in the adequacy era: Heterogeneous causal effects and sensitivity." *Education Finance and Policy*, *14*(1): 31-60. https://doi.org/10.1162/edfp_a_00236

Carneiro, Pedro, and Rita Ginja. 2014. "Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start." *American Economic Journal: Economic Policy*, 6(4), 135–173. https://doi.org/10.1257/pol.6.4.135

Cho, Hyunkuk, Paul Glewwe, and Melissa Whitler. 2012. "Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools." *Economics of Education Review*, 31(3): 77-95. https://doi.org/10.1016/j.econedurev.2012.01.004

Davezies, Laurent, and Manon Garrouste. 2020. "More harm than good?: Sorting effects in a compensatory education program." *Journal of Human Resources*, 55(1): 240-277. https://doi.org/10.3368/jhr.55.1.0416-7839R1

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110 (9): 2964–96. https://doi.org/10.1257/aer.20181169

De Haan, Monique. 2017. "The Effect of Additional Funds for Low-ability Pupils: A Non-parametric Bounds Analysis." *The Economic Journal*, *127*(599): 177-198. https://doi.org/10.1111/ecoj.12221

Fairlie, Robert W., and Alexandra M. Resch. 2002. "Is there "white flight" into private schools? Evidence from the National Educational Longitudinal Survey." *Review of Economics and Statistics*, 84(1): 21-33. https://doi.org/10.1162/003465302317331892

Franck, Emilie, and Ides Nicaise. 2022. "The effectiveness of equity funding policies in schools in Europe and North America: A systematic literature review." *Issues in Educational Research*, 32(2): 494–512. https://doi/10.3316/informit.579200243997465

Fritsch, Rosangela, and Carlinda Leite. 2020. "Territórios educativos de intervenção prioritária: um inventário do conhecimento divulgado em periódicos." *Revista Conhecimento Online*, 3: 145-171. https://doi.org/10.25112/rco.v3i0.2301

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*, 225(2): 254-277. https://doi.org/10.1016/j.jeconom.2021.03.014

Gordon, Nora. 2004. "Do federal grants boost school spending? Evidence from Title I." *Journal of Public Economics*, 88(9-10): 1771-1792. https://doi.org/10.1016/j.jpubeco.2003.09.002

Hægeland, Torbjørn, Oddbjørn Raaum, and Kjell G. Salvanes. (2012). "Pennies from heaven? Using exogenous tax variation to identify effects of school resources on pupil achievement." *Economics of Education Review*, 31(5): 601-614. https://doi.org/10.1016/j.econedurev.2012.03.004

Hanushek, Eric Alan. 2003. "The failure of input-based schooling policies." The *Economic Journal*, 113 (485): F64-F98. https://doi.org/10.1111/1468-0297.00099

Hyman, Joshua. 2017. "Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment." *American Economic Journal: Economic Policy*, 9(4): 256–80. https://doi.org/10.1257/pol.20150249

Jackson, C. Kirabo, Rucker C. Johnson and Claudia Persico. 2016. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance

Reforms." *The Quarterly Journal of Economics*, 131 (1): 157–218. https://doi.org/10.1093/qje/qjv036

Jackson, C. Kirabo. 2020. "Does school spending matter? The new literature on an old question." In L. Tach, R. Dunifon, & D. L. Miller (Eds.), *Confronting inequality: How policies and practices shape children's opportunities* (pp. 165–186). American Psychological Association. https://doi.org/10.1037/0000187-008

Jackson, C. Kirabo, and Claire L. Mackevicius. 2024. "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States." *American Economic Journal: Applied Economics*, 16 (1): 412–46. https://doi.org/10.1257/app.20220279

Johnson, Rucker C., and C. Kirabo Jackson. 2019. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." *American Economic Journal: Economic Policy*, 11 (4): 310–49. https://doi/10.1257/pol.20180510

Krueger, Alan. 2003. "Economic Considerations and Class Size." *The Economic Journal*, 113 (485): F34–F63. https://doi.org/10.1111/1468-0297.00098

Leuven, Edwin, Mikael Lindahl, Hessel Oosterbeek, and Dinand Webbink. 2007. "The effect of extra funding for disadvantaged pupils on achievement." *The Review of Economics and Statistics*, *89*(4): 721-736. https://doi.org/10.1162/rest.89.4.721

Machin, Stephen, Sandra McNally, and Costas Meghir. 2010. "Resources and Standards in Urban Schools." *Journal of Human Capital*, 4(4): 365-393. https://doi.org/10.1086/658634

Matsudaira, Jordan D., Adrienne Hosek, and Elias Walsh. 2012. "An integrated assessment of the effects of Title I on school behavior, resources, and student achievement." *Economics of Education Review*, 31(3): 1-14. https://doi.org/10.1016/j.econedurev.2012.01.002

Rios-Avila, Fernando, Pedro Sant'Anna, and Brantly Callaway. 2023. "CSDID: Stata module for the estimation of Difference-in-Difference models with multiple time periods." In *Statistical Software Components*. Boston College Department of Economics.

Schelling, Thomas. 1969. "Models of segregation." *The American economic review*, *59*(2): 488-493. https://www.jstor.org/stable/1823701

———. 1971. "Dynamic models of segregation." *Journal of mathematical sociology*, *1*(2): 143-186. https://doi.org/10.1080/0022250X.1971.9989794

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of econometrics*, 225(2): 175-199. https://doi.org/10.1016/j.jeconom.2020.09.006

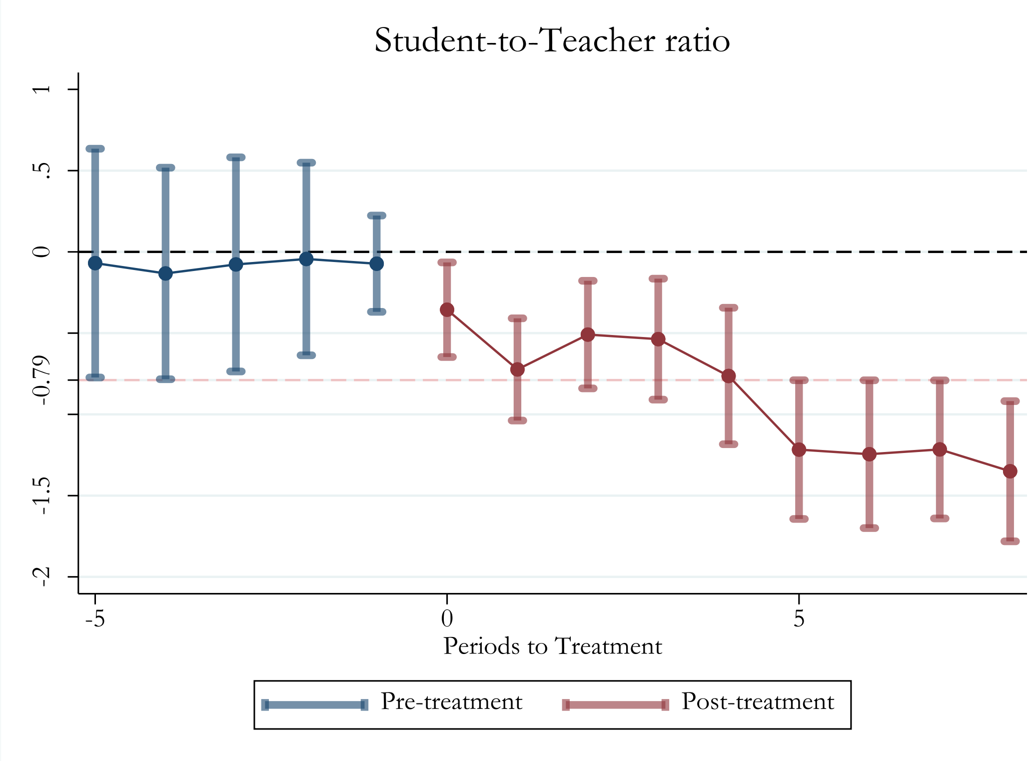
Sun, Liyang. 2022. "EVENTSTUDYINTERACT: Stata module to implement the interaction weighted estimator for an event study." In *Statistical Software Components*. Boston College Department of Economics.

Van der Klaauw, Wilbert. 2008. "Breaking the link between poverty and low student achievement: An evaluation of Title I." *Journal of Econometrics*, 142(2): 731-756. https://doi.org/10.1016/j.jeconom.2007.05.007

Van Ewijk, Reyn, and Peter Sleegers. 2010. “The effect of peer socioeconomic status on student achievement: A meta-analysis.” *Educational research review*, *5*(2): 134-150. https://doi.org/10.1016/j.edurev.2010.02.001

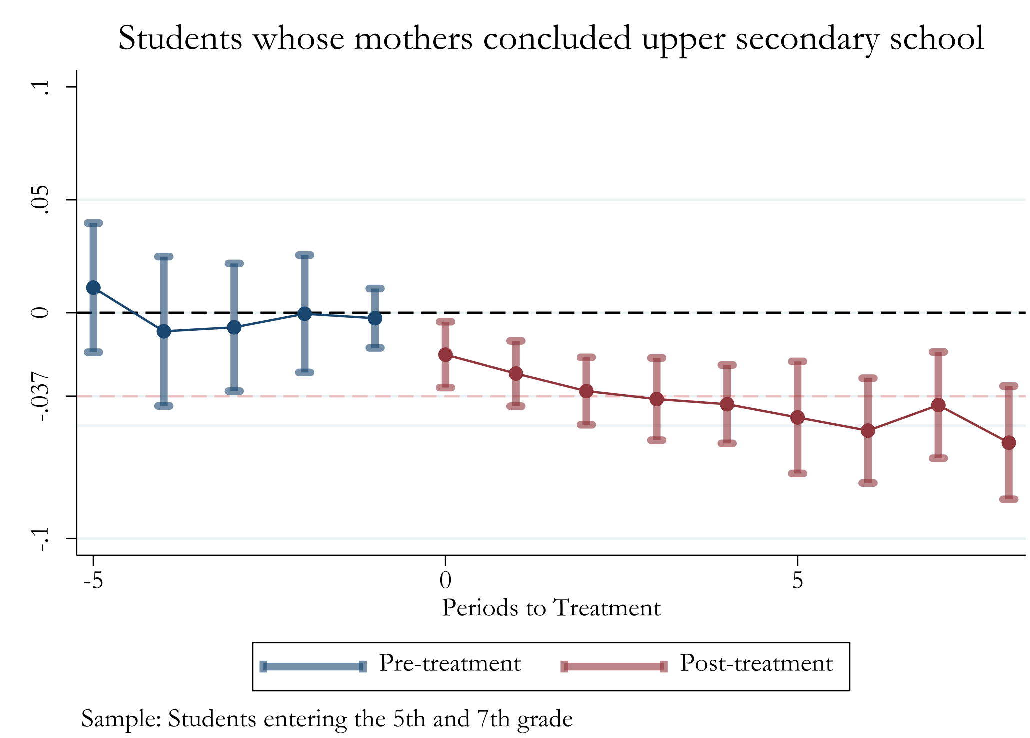
Figures

Figure 1. Additional resources: Student-to-Teacher ratio



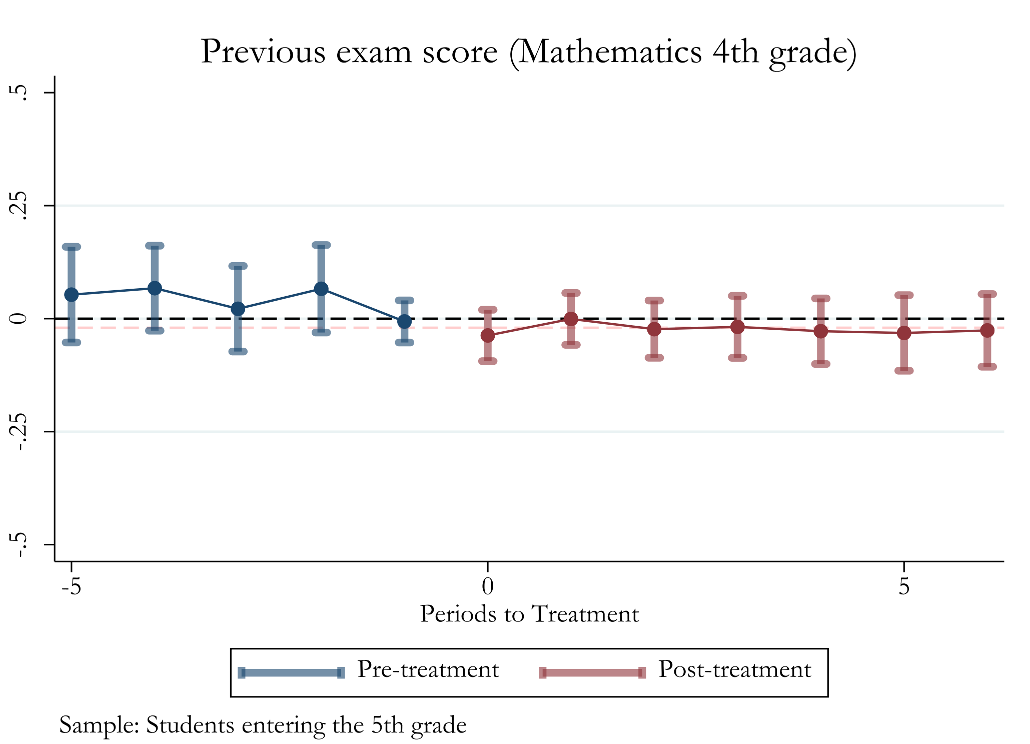
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest, the Student-to-Teacher ratio, is measured at school level. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Standard errors are clustered at school level. The bars represent 95% confidence intervals.

Figure 2. Composition of students: Students whose mothers concluded upper secondary school



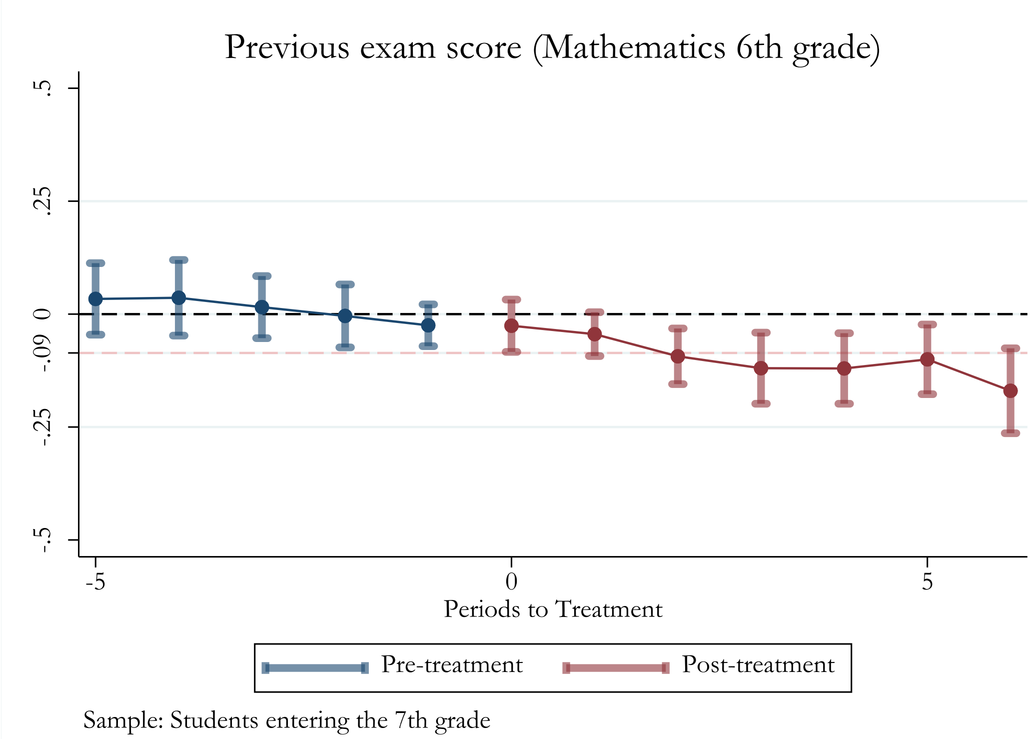
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a dummy variable which takes the value of 1 if a student's mother has concluded upper secondary school. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure 3. Composition of students: Previous exam scores (Mathematics 4th grade)



Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Mathematics 4th grade exam score obtained in the year before entering the 5th grade. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure 4. Composition of students: Previous exam scores (Mathematics 6th grade)



Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Mathematics 6th grade exam score obtained in the year prior to entering the 7th grade. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Tables

Table 1. Student achievement – Exam scores

Sub-sample:	**Mother has not concluded secondary school**		**Receiving social support – Type A**	
Dep. Var.:	**Mathematics exam score**	**Portuguese exam score**	**Mathematics exam score**	**Portuguese exam score**
	(1)	**(2)**	**(3)**	**(4)**
Panel A: 9th grade (0-100 scale)				
TEIP schools * Post-Treatment	-0.678	-0.421	-0.214	-1.275**
	(0.71)	(0.51)	(0.81)	(0.59)
Pre-trend test	[0.25]	[0.24]	[0.57]	[0.19]
N of clusters	1 182	1 182	1 157	1 157
N	485 683	484 975	162 226	161 465
Panel B: 6th grade (1-5 scale)				
TEIP schools * Post-Treatment	-0.021	0.004	-0.001	0.014
	(0.02)	(0.02)	(0.02)	(0.02)
Pre-trend test	[0.60]	[0.16]	[0.24]	[0.09]
N of clusters	886	886	879	879
N	473 171	473 850	199 933	200 257
Panel C: 4th grade (1-5 scale)				
TEIP schools * Post-Treatment	0.052*	0.001	0.017	-0.010
	(0.03)	(0.02)	(0.04)	(0.03)
Pre-trend test	[0.56]	[0.38]	[0.57]	[0.80]
N of clusters	914	914	889	889
N	427 781	428 513	139 546	139 785

Notes: Our variable of interest, *TEIP schools* * Post-Treatment, is a dummy variable with a value of 1 for students attending a *TEIP* school after it adheres to the program. The pre-trend test value in square brackets corresponds to the p-value of the test of the null hypothesis that all pre-treatment ATTs, by group and year, are equal to zero. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Clustered standard errors at school cluster level are presented in parentheses. Significance level at which the null hypothesis is rejected: *** 1%, ** 5%, * 10%.

Table 2. Student achievement – Percentage of passing scores in exams

Sub-sample:	**Mother has not concluded secondary school**	**Receiving social support – Type A**
Dep. Var.:	**Percentage of passing scores in exams**	
	(1)	**(2)**
Panel A: Baseline		
TEIP schools * Post-Treatment	-0.004	-0.001
	(0.01)	(0.01)
Pre-trend test	[0.12]	[0.04]
Number of Clusters	1 267	1 228
N	1 282 309	456 503
Panel B: By *TEIP* phase		
TEIP 2/3 schools * Post-Treatment	-0.004	-0.002
	(0.01)	(0.01)
TEIP 4 schools * Post-Treatment	-0.001	0.007
	(0.01)	(0.02)
Panel C: By educational stage		
Lower Secondary (9th grade): *TEIP schools* * Post-Treatment	-0.010	-0.010
	(0.01)	(0.02)
2nd cycle (6th grade): *TEIP schools* * Post-Treatment	-0.003	0.005
	(0.01)	(0.01)
1st cycle (4th grade): *TEIP schools* * Post-Treatment	-0.004	-0.006
	(0.01)	(0.02)

Notes: Our variable of interest, *TEIP schools* * Post-Treatment, is a dummy variable with a value of 1 for students attending a *TEIP* school after it has adhered to the program. The pre-trend test value in square brackets corresponds to the p-value of the test of the null hypothesis that all pre-treatment ATTs, by group and year, are equal to zero. Grade fixed-effects are present in Panel A and Panel B. We only analyze students who did both Mathematics and Portuguese exams. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Clustered standard errors at school cluster level are presented in parentheses. Significance level at which the null hypothesis is rejected: *** 1%, ** 5%, * 10%.

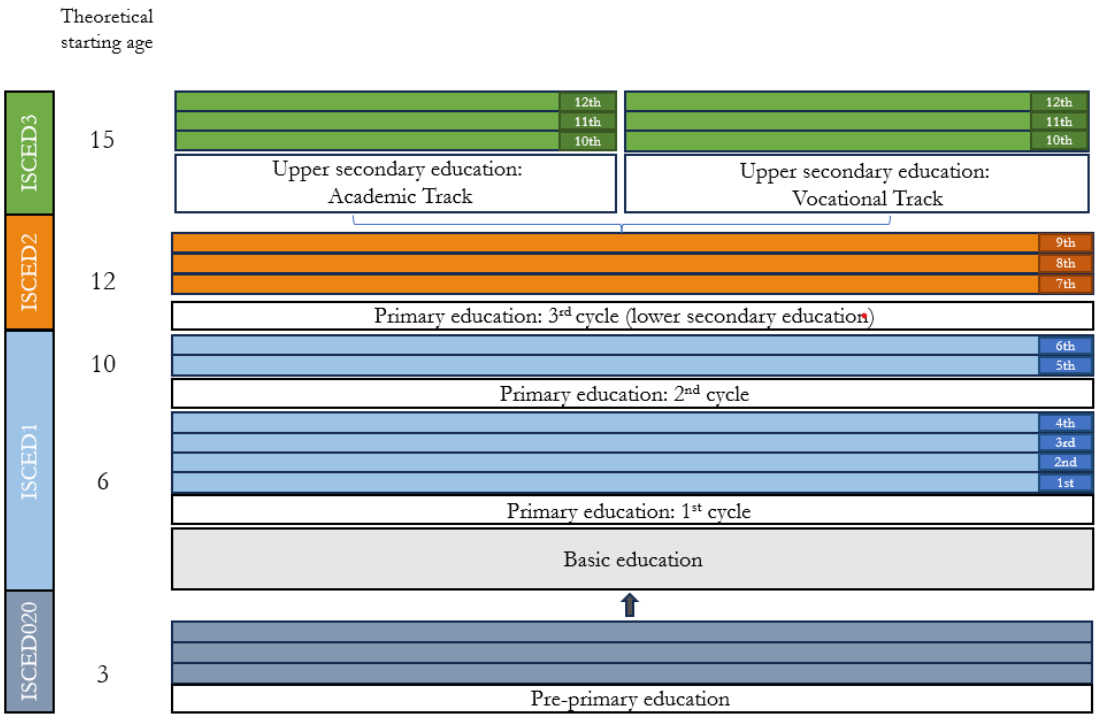
Table 3. Student achievement – Internal sources of evaluation

Sub-sample:	Mother has not concluded secondary school	Receiving social support – Type A
	(1)	(2)
Panel A: Successful progress in lower secondary education		
TEIP schools * Post-Treatment	0.002	0.004
	(0.01)	(0.01)
Pre-trend test	[0.39]	[0.87]
N of clusters	1 093	1 087
N	435 406	186 811
Panel B: 7th and 8th grade without retention		
TEIP schools * Post-Treatment	0.033***	0.025*
	(0.01)	(0.01)
Pre-trend test	[0.82]	[0.21]
TEIP 2/3 schools * Post-Treatment	0.041***	0.031**
	(0.01)	(0.02)
TEIP 4 schools * Post-Treatment	-0.004	0.002
	(0.02)	(0.02)
N of clusters	1 093	1 087
N	435 406	186 811
Panel C: 5th and 6th grade without retention		
TEIP schools * Post-Treatment	-0.001	-0.019
	(0.01)	(0.01)
Pre-trend test	[0.31]	[0.02]
N of clusters	849	848
N	430 918	202076

Notes: Our variable of interest, *TEIP schools* * Post-Treatment, is a dummy variable with a value of 1 for students attending a *TEIP* school after it has adhered to the program. The pre-trend test value in square brackets corresponds to the p-value of the test of the null hypothesis that all pre-treatment ATTs, by group and year, are equal to zero. Heterogeneity across different *TEIP* phases is not presented in Panels A and C as the estimates are non-significant and of similar magnitude across the panel. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Clustered standard errors at school cluster level are presented in parentheses. Significance level at which the null hypothesis is rejected: *** 1%, ** 5%, * 10%.

Appendix

Figure A1. Portuguese educational system



Notes: In basic education, students can opt to enroll in the general track, as shown in the figure, or other specific tracks. As the latter represents less than 4% of the students in basic education, it is not represented in the figure.

Figure A2. Number of *TEIP* school clusters over time

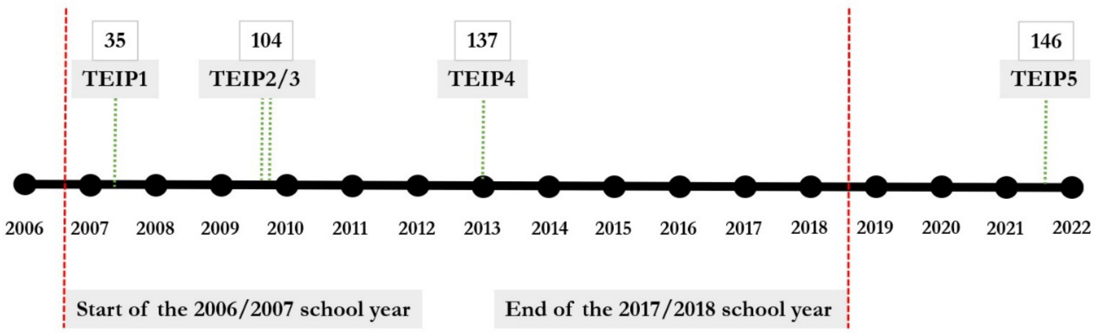
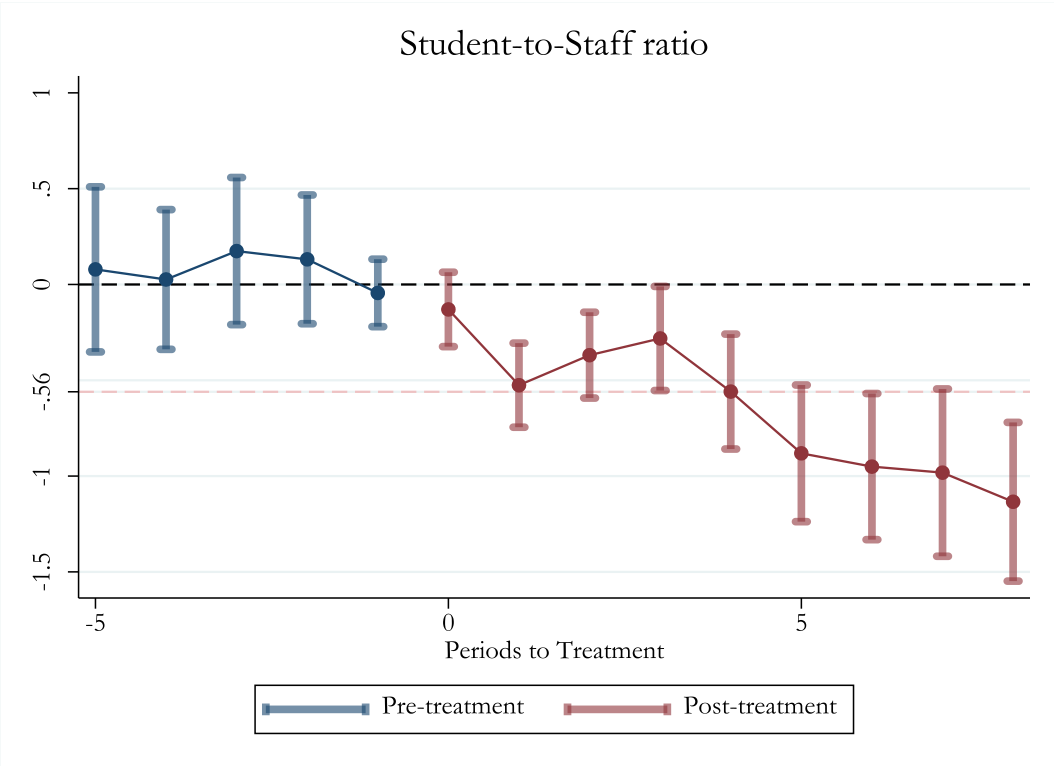
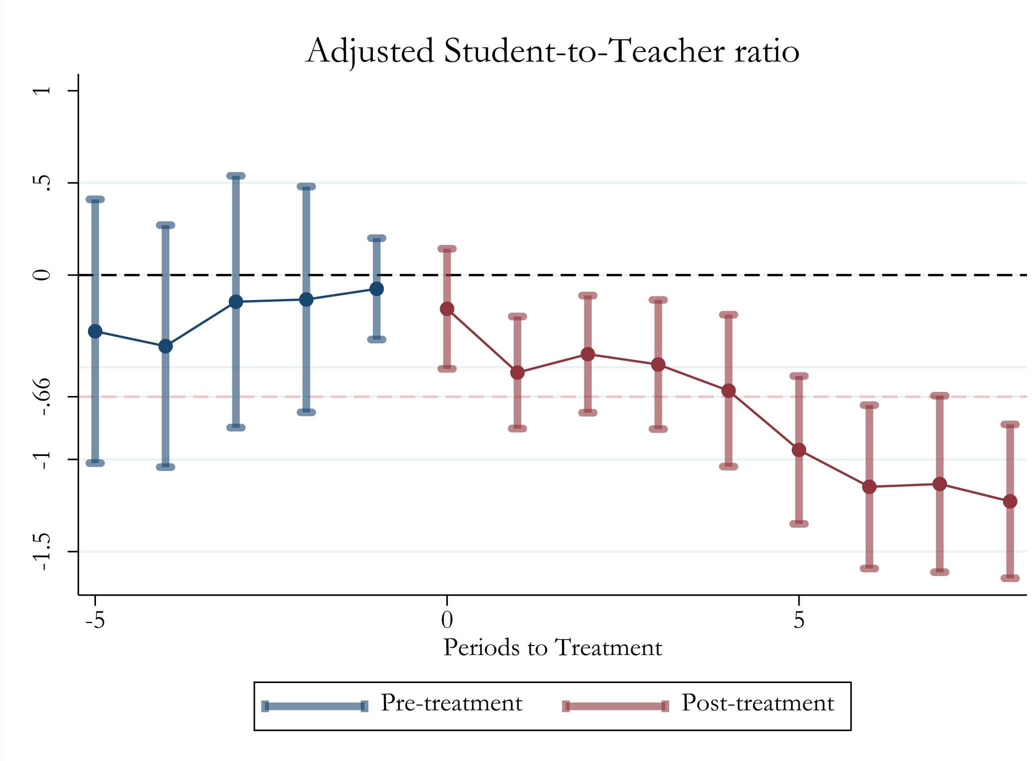


Figure A3. Additional resources: Student-to-Staff ratio



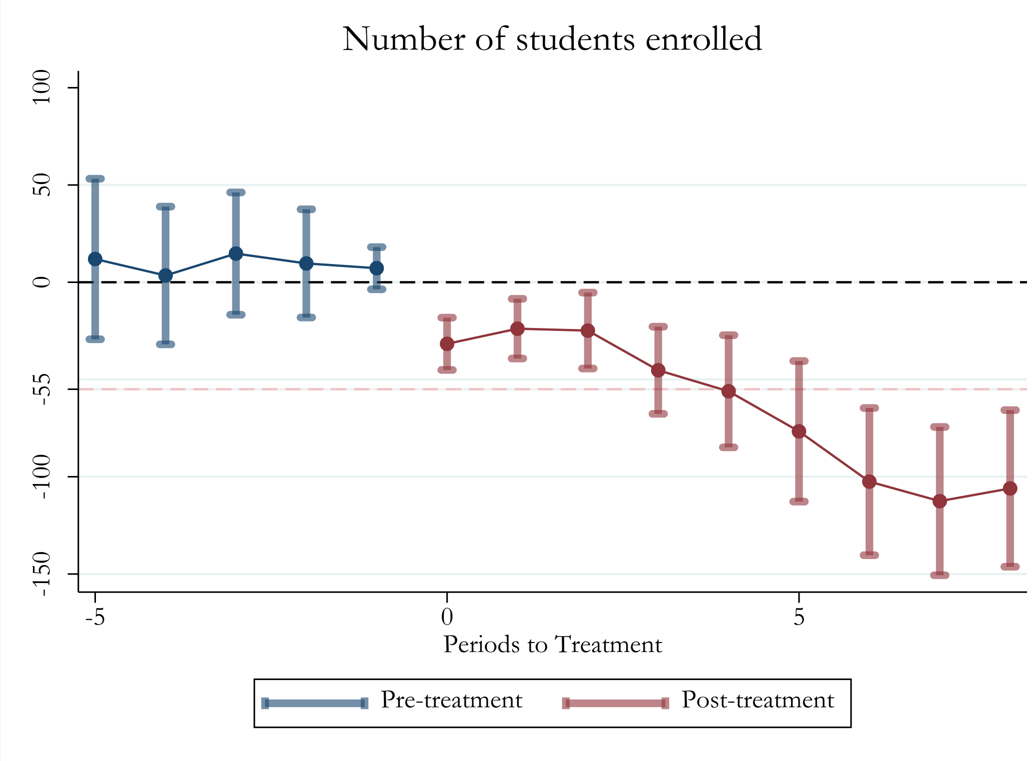
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest, the Student-to-Staff ratio, is measured at school level. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A4. Additional resources: Adjusted Student-to-Teacher ratio



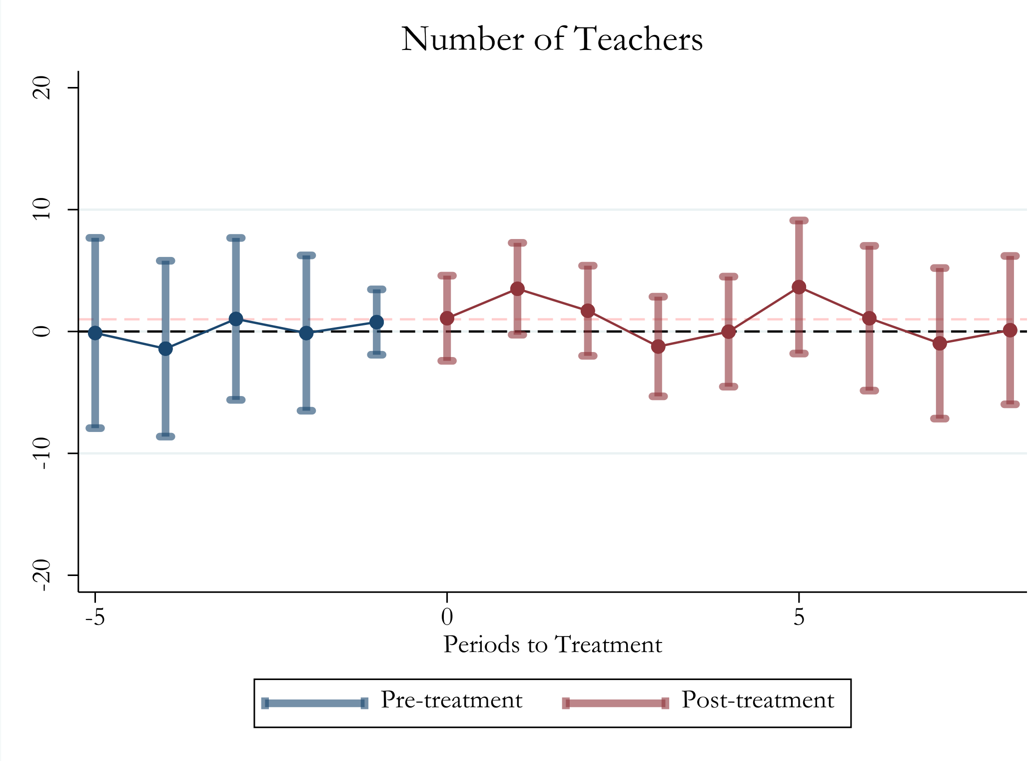
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest, the Adjusted Student-to-Teacher ratio, is measured at school level. It takes into consideration teachers that are assigned to more than one school The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Standard errors are clustered at school level. The bars represent 95% confidence intervals.

Figure A5 Additional resources: Number of students



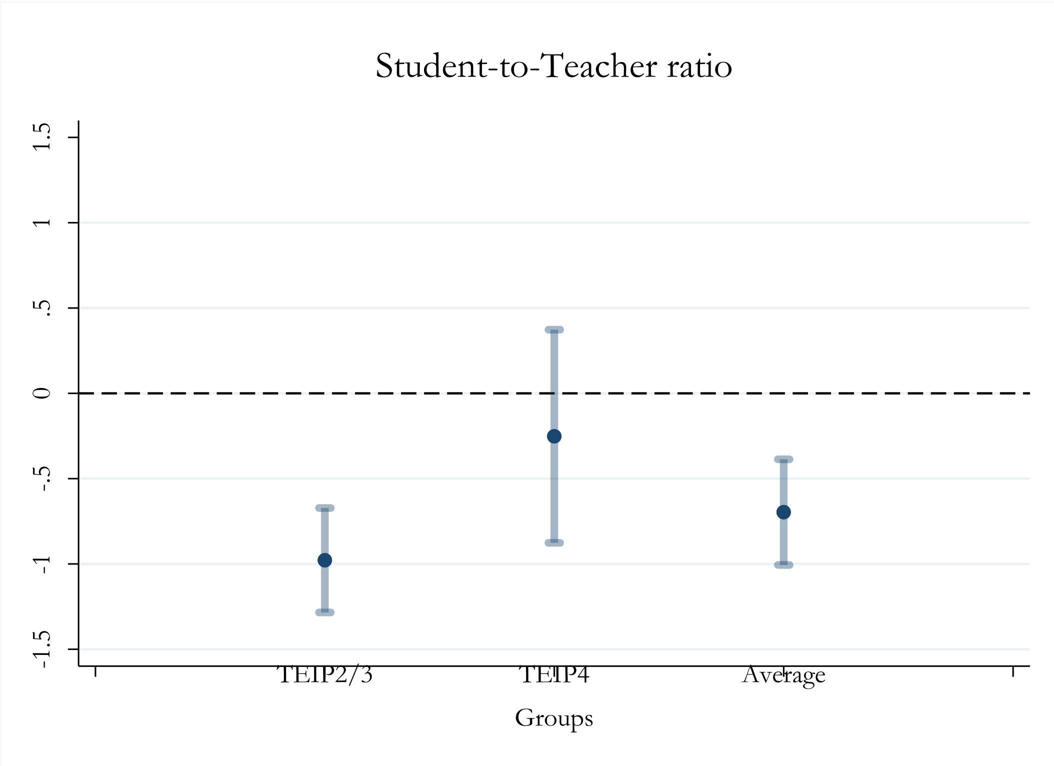
Notes: This graph shows the ATT estimates across TEIP phase groups using the Callaway and Sant'Anna (2021) method. Our outcome of interest, the Student-to-Teacher ratio, is measured at school level. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A6. Additional resources: Number of teachers



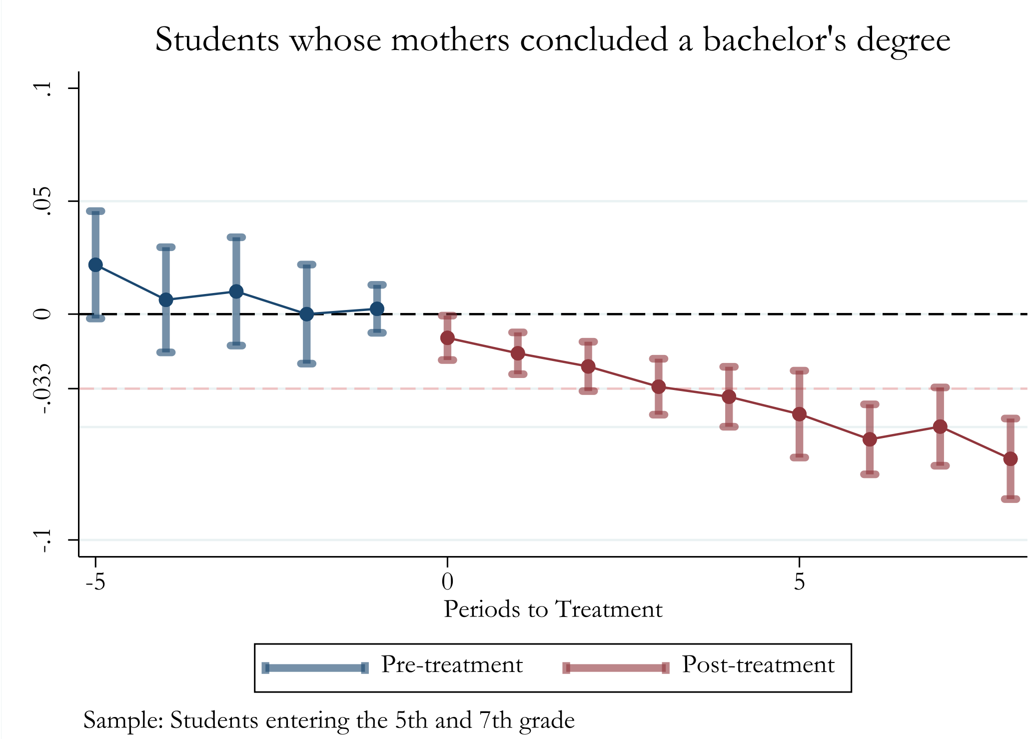
Notes: This graph shows the ATT estimates across TEIP phase groups using the Callaway and Sant'Anna (2021) method. Our outcome of interest, the Student-to-Teacher ratio, is measured at school level. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A7. Additional resources: Student-to-Teacher ratio by TEIP phase



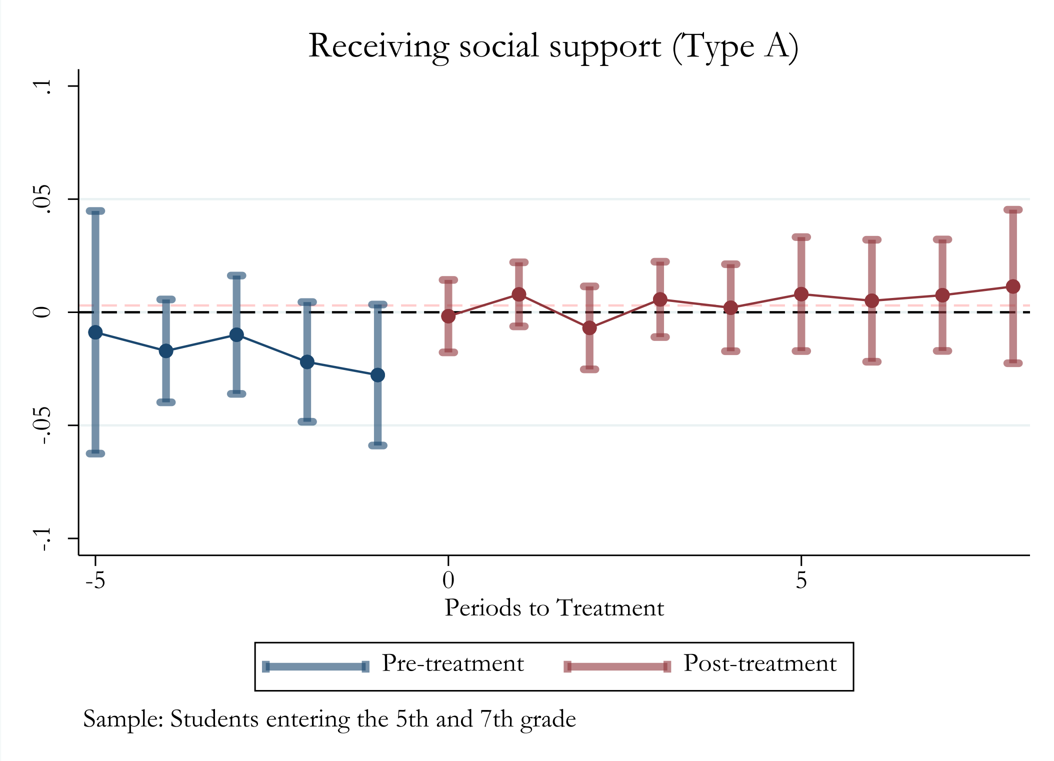
Notes: This graph shows the ATT estimates across TEIP phase groups using the Callaway and Sant'Anna (2021) method. Our outcome of interest, the Student-to-Teacher ratio, is measured at school level. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A8. Composition of students: Students whose mothers concluded a bachelor's degree



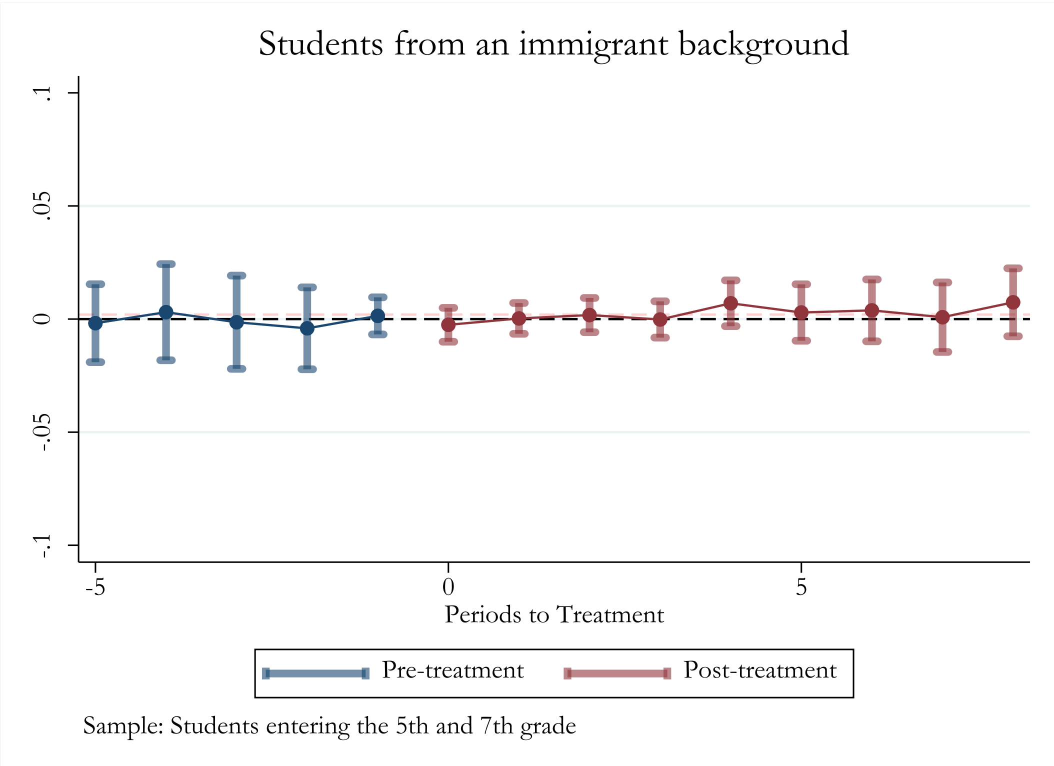
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a dummy variable which takes the value of 1 if a student's mother has a bachelor's degree. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A9. Composition of students: Students receiving social support (Type A)



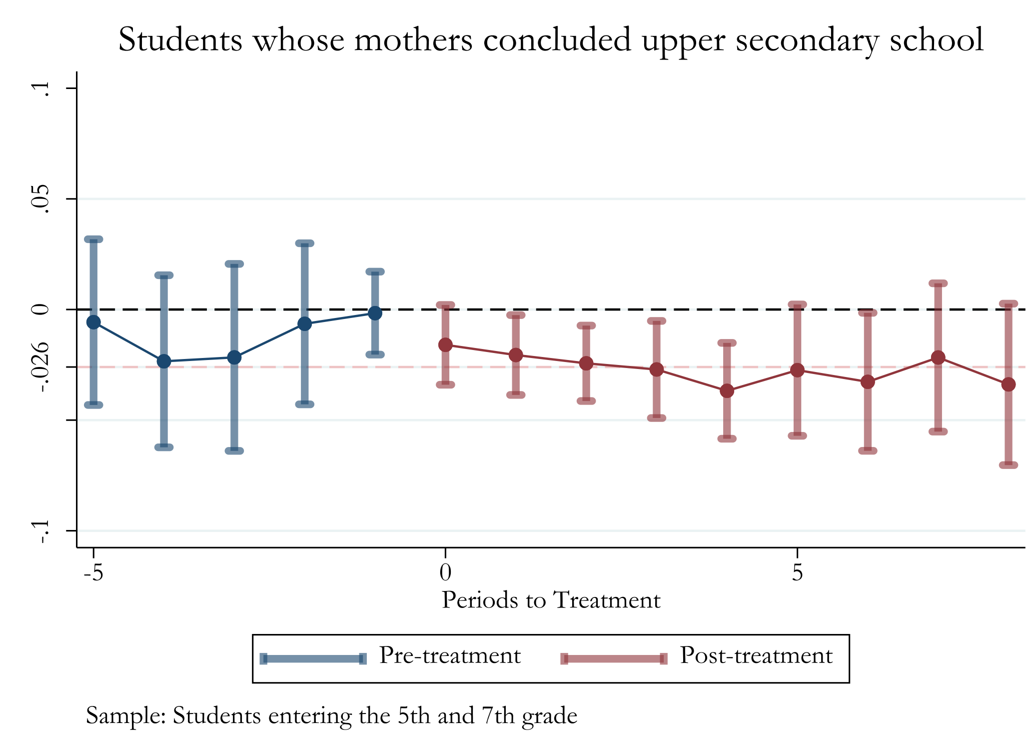
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a dummy variable, which takes the value of 1 if a student receives the highest level of social support from the government, Type A. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A10. Composition of students: Students from an immigrant background



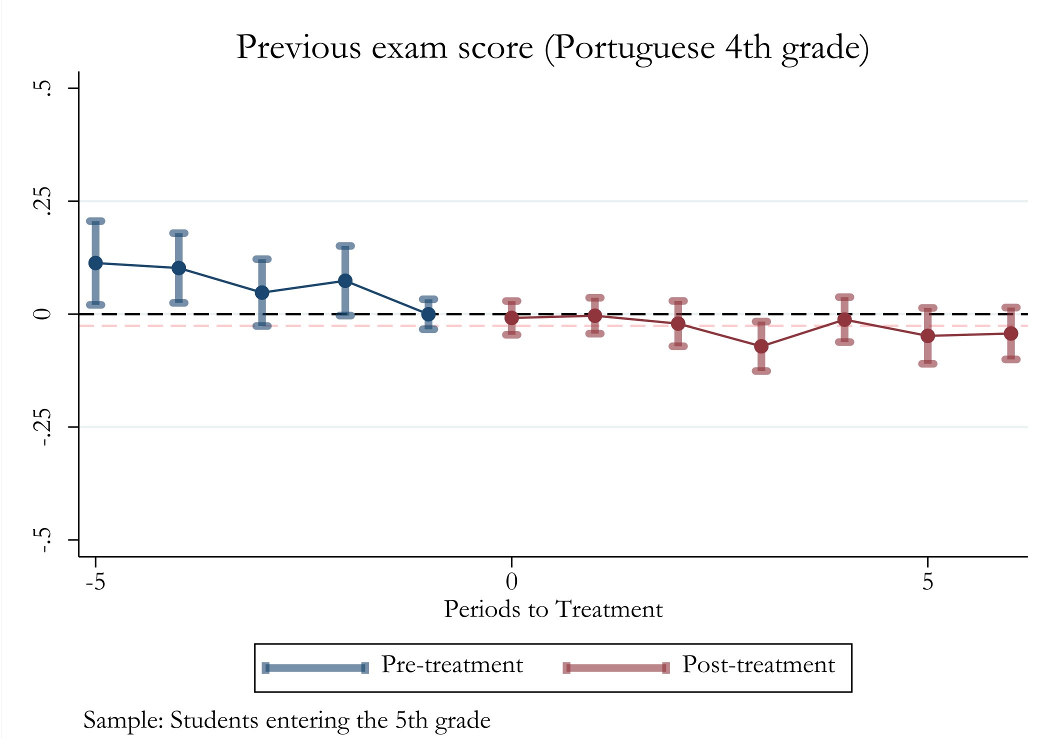
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a dummy variable that takes the value of 1 if a student is from an immigrant background, defined as not having Portuguese nationality or having at least one parent who is not Portuguese. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A11. Composition of students: Students whose mothers concluded upper secondary school



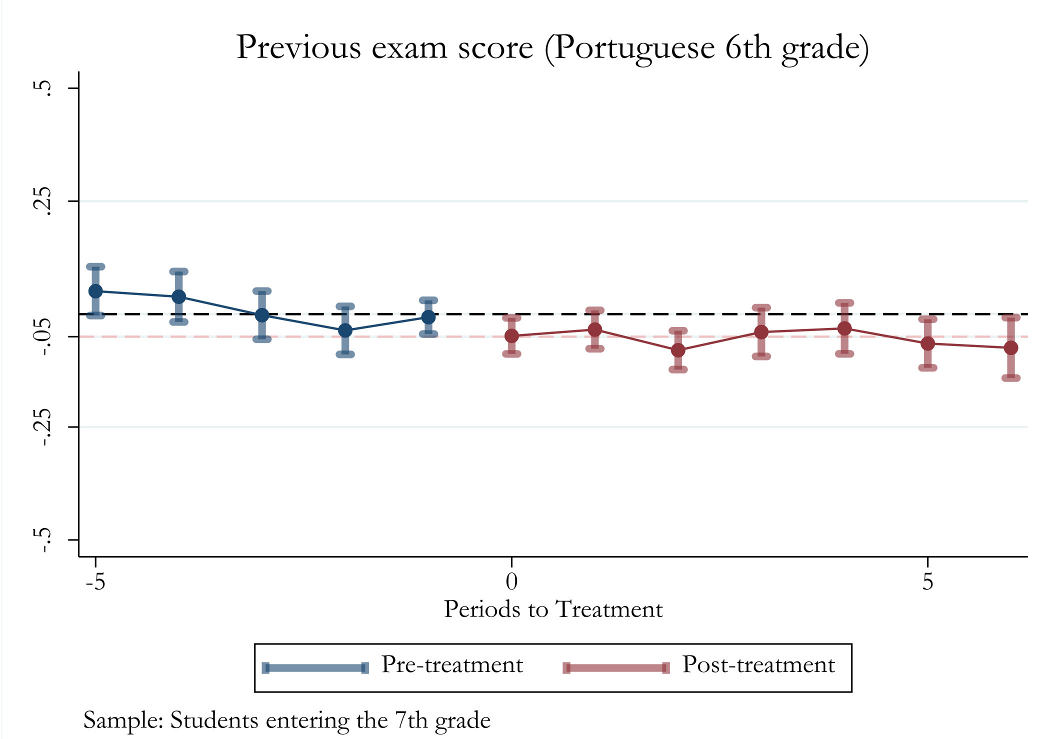
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a dummy variable which takes the value of 1 if a student's mother has concluded upper secondary school. Year × municipality fixed-effects are introduced. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A12. Composition of students: Previous exam scores (Portuguese 4th grade)



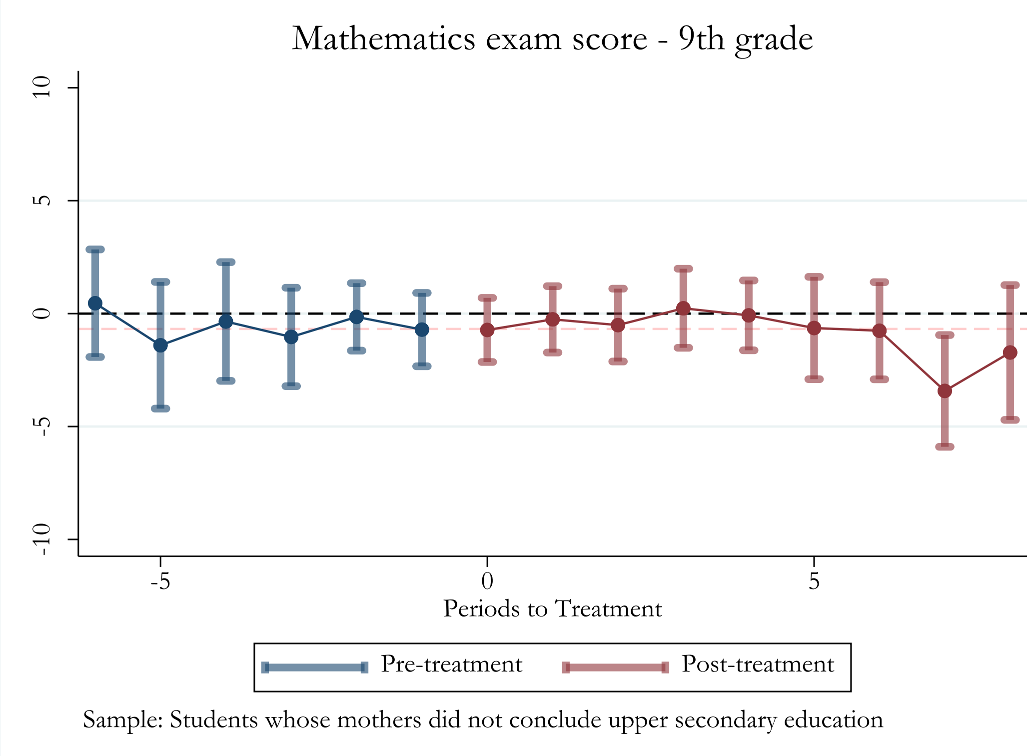
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Portuguese 4th grade exam score obtained the year before entering the 5th grade. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A13. Composition of students: Previous exam scores (Portuguese 6th grade)



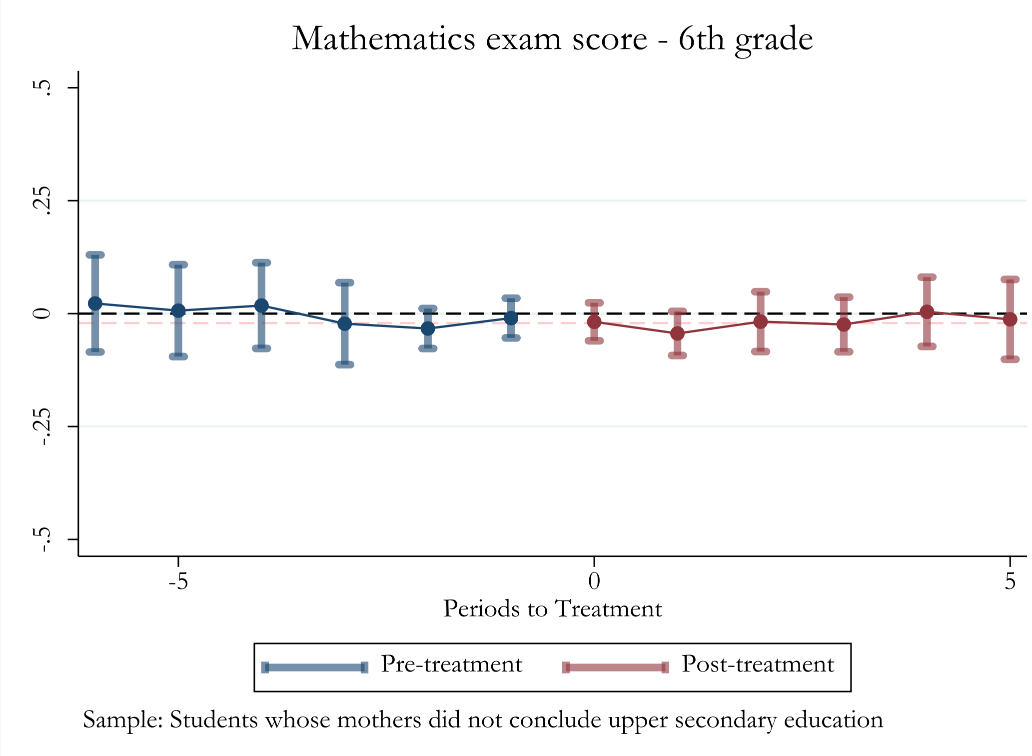
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Portuguese 6th grade exam score obtained the year before entering the 7th grade. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A14. Student achievement – Mathematics 9th grade exam score



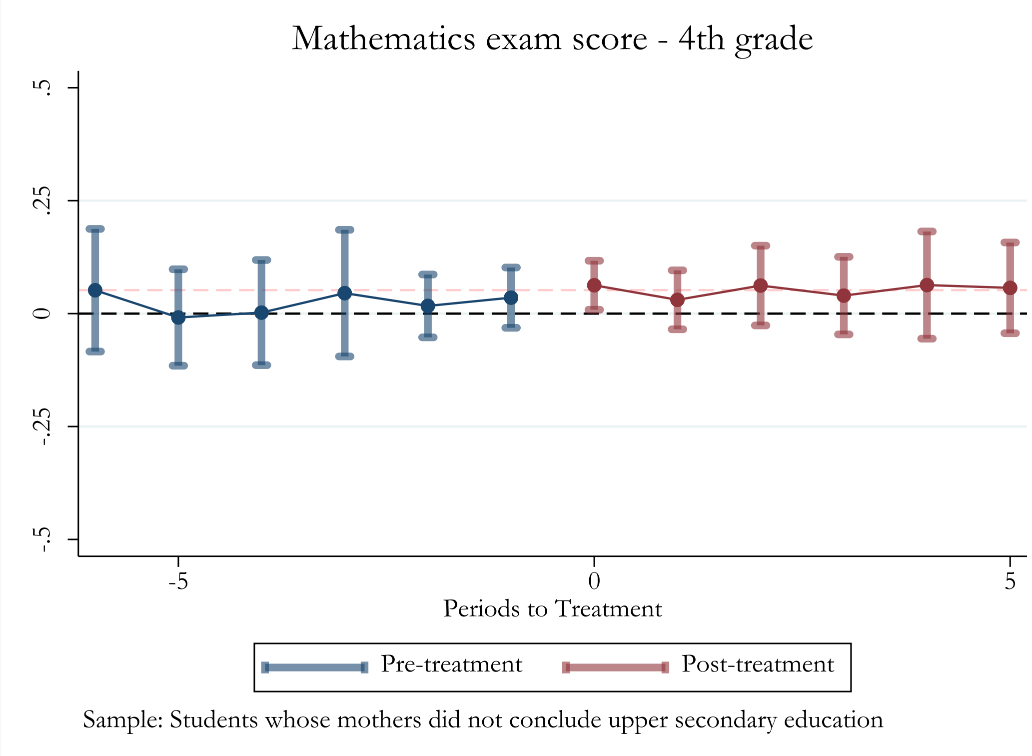
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Mathematics 9th grade exam score. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A15. Student achievement – Mathematics 6th grade exam score



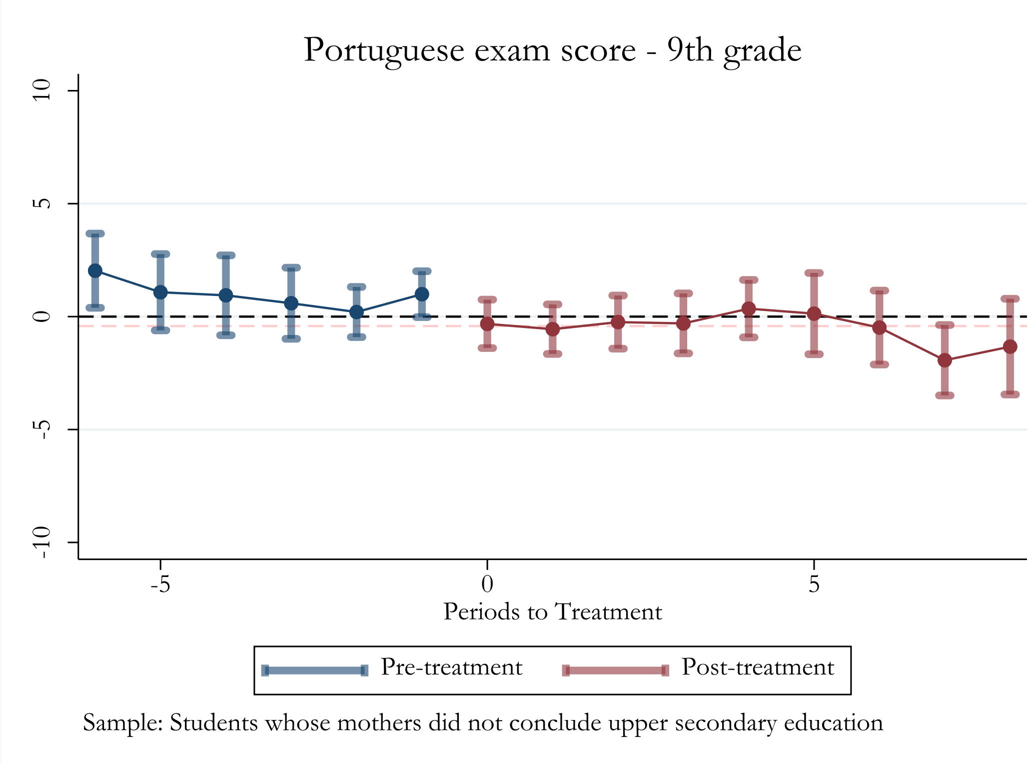
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Mathematics 6th grade exam score. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2014/2015 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A16. Student achievement – Mathematics 4th grade exam score



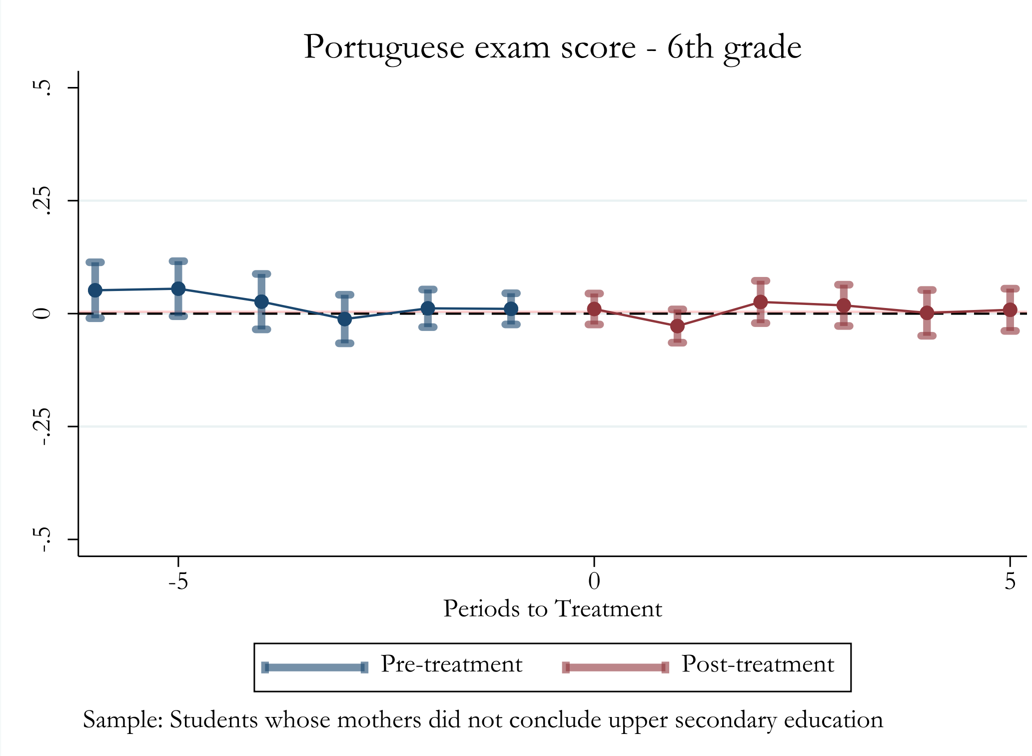
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Mathematics 4th grade exam score. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2014/2015 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A17. Student achievement – Portuguese 9th grade exam score



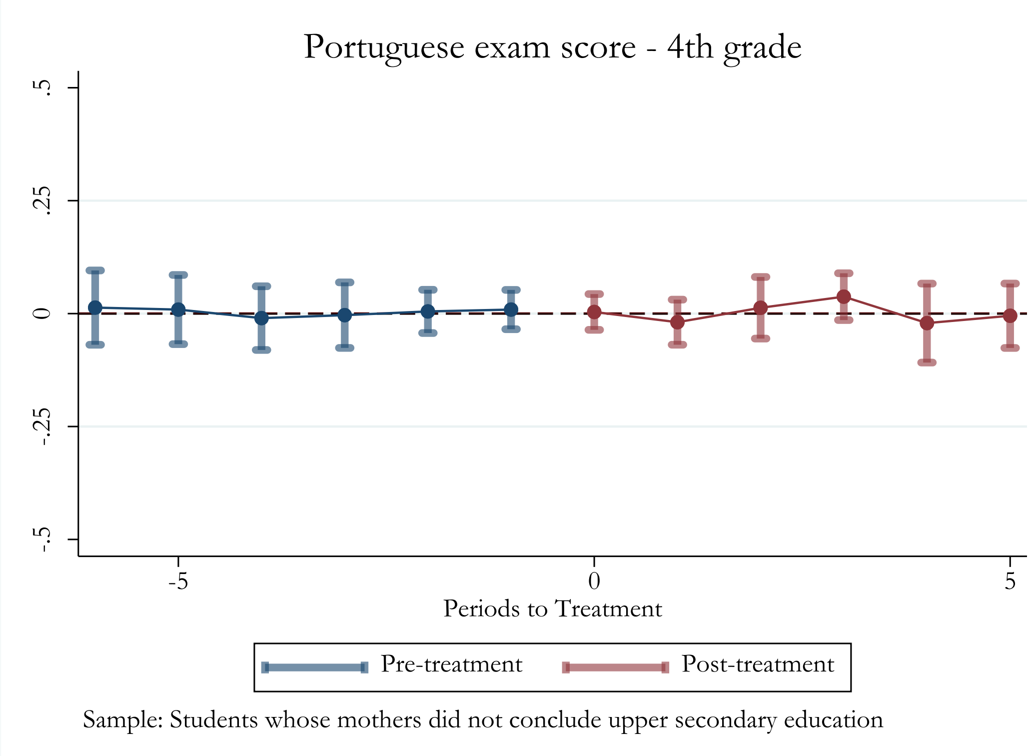
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Portuguese 9th grade exam score. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A18. Student achievement – Portuguese 6th grade exam score



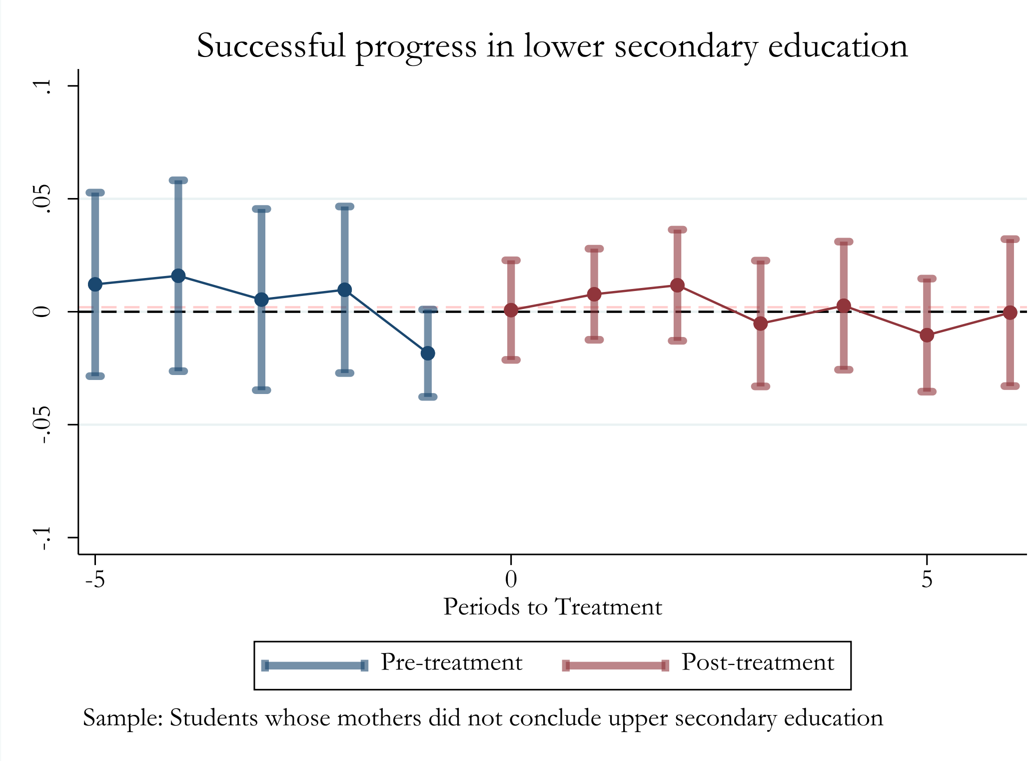
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Portuguese 6th grade exam score. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2014/2015 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A19. Student achievement – Portuguese 4th grade exam score



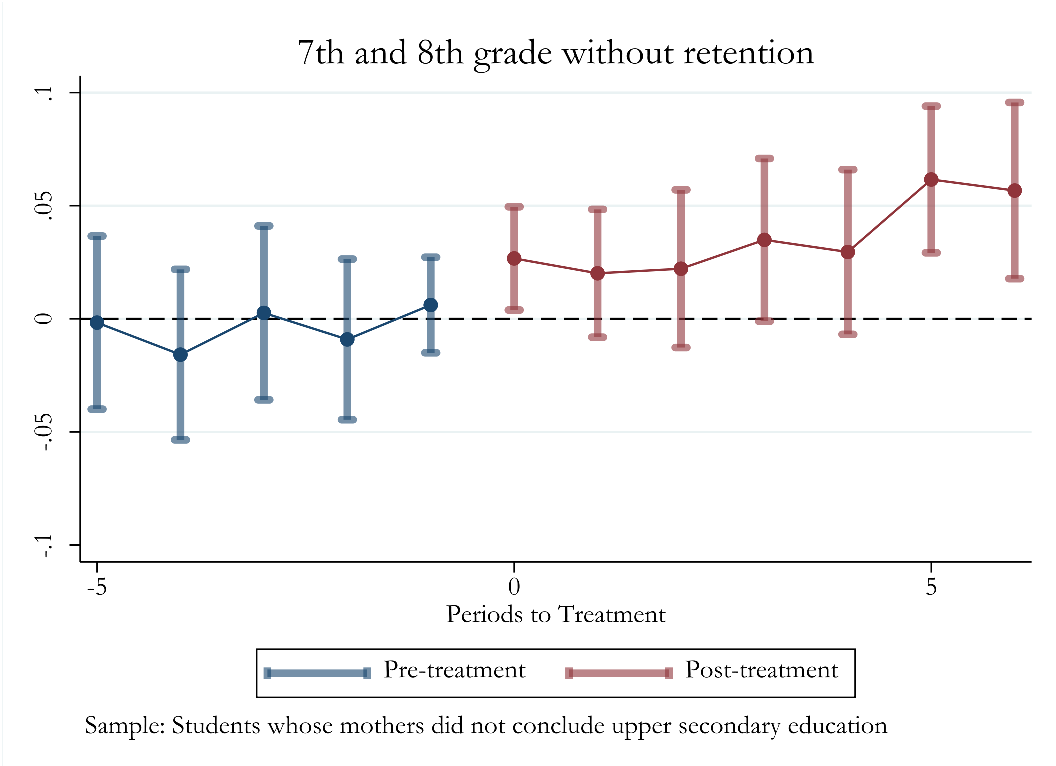
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is the Portuguese 4th grade exam score. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2014/2015 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A20. Student achievement – Successful progress in lower secondary education



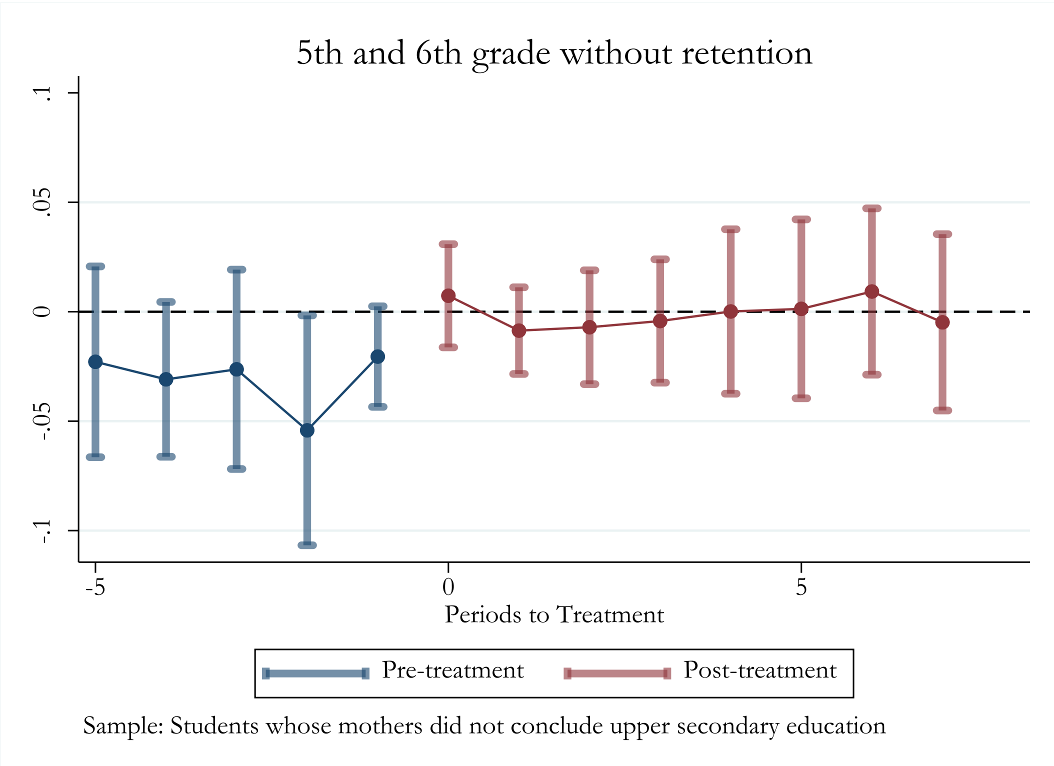
Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a binary indicator that takes the value of one if a student passes both the 7th and 8th grade without retention and has a passing score in the Mathematics and Portuguese 9th grade exams. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A21. Student achievement – 7th and 8th grade without retention



Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a binary indicator that takes the value of one if a student passes both the 7th and 8th grade without retention. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Figure A22. Student achievement – 5th and 6th grade without retention



Notes: This graph shows the event-study estimates using the Callaway and Sant'Anna (2021) method. Our outcome of interest is a binary indicator that takes the value of one if a student passes both the 7th and 8th grade without retention. The base period is the year before entering the *TEIP* program (g-1), both for pre- and post-treatment (long2 option in the csdid command). The red dashed line is the ATT value across all periods and groups. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Standard errors are clustered at school cluster level. The bars represent 95% confidence intervals.

Table A1. Descriptive Statistics: Entire sample

Panel A: Entire Sample	N	Mean	SD	Range
Percentage of passing scores in 9th grade exams	906 575	0.59	0.39	0 – 1
Mathematics 9th grade exam score	913 475	46.96	24.70	0 – 100
Portuguese 9th grade exam score	910 238	56.13	16.10	0 – 100
Percentage of passing scores in 6th grade exams	825 931	0.71	0.38	0 – 1
Mathematics 6th grade exam score	833 017	2.84	0.98	1 – 5
Portuguese 6th grade exam score	833 056	3.10	0.78	1 – 1
Percentage of passing scores in 4th grade exams	796 361	0.79	0.35	0 – 1
Mathematics 4th grade exam score	799 674	3.21	0.98	0 – 100
Portuguese 4th grade exam score	800 473	3.26	0.85	0 – 100
Successful progress in lower secondary education	808 191	0.39	0.49	0 – 1
7th and 8th grade without retentions	808 191	0.75	0.43	0 – 1
5th and 6th grade without retentions	843 564	0.72	0.45	0 – 1
Mother has concluded secondary school	9 465 323	0.41	0.49	0 – 1
Mother has a bachelor's degree	9 465 323	0.18	0.38	0 – 1
Father has concluded secondary school	8 980 248	0.33	0.47	0 – 1
Father has a bachelor's degree	8 980 248	0.12	0.33	0 – 1
Receiving social support - Type A	10 244 148	0.23	0.42	0 – 1
Student from an immigrant background	10 135 716	0.09	0.29	0 – 1

Table A2. Descriptive Statistics: Students whose mothers did not conclude secondary education

***TEIP* vs *Non-TEIP* (Pre-2009)**						
Panel A: Mother has not concluded secondary education	*Non-TEIP*			*TEIP*		
	N	Mean	SD	N	Mean	SD
Percentage of passing scores in 9th grade exams	128 117	0.59	0.35	15 197	0.53	0.35
Mathematics 9th grade exam score	128 724	43.77	22.45	15 312	39.81	21.53
Portuguese 9th grade exam score	128 651	57.42	14.48	15 282	55.73	14.69
Percentage of passing scores in 6th grade exams	151 153	0.78	0.33	25 707	0.72	0.36
Mathematics 6th grade exam score	152 478	2.85	0.83	26 009	2.71	0.82
Portuguese 6th grade exam score	152 919	3.11	0.67	26 163	3.03	0.67
Percentage of passing scores in 4th grade exams	148 729	0.87	0.28	21 931	0.82	0.32
Mathematics 4th grade exam score	149 469	3.37	0.90	22 087	3.21	0.92
Portuguese 4th grade exam score	149 695	3.26	0.74	22 150	3.16	0.77
Successful progress in lower secondary education	97 627	0.29	0.45	13 233	0.24	0.42
7th and 8th grade without retentions	97 627	0.69	0.46	13 233	0.65	0.48
5th and 6th grade without retentions	92 343	0.69	0.46	16 381	0.67	0.47
Receiving social support - Type A	931 737	0.25	0.43	144 512	0.31	0.46
Panel B: Receiving social support - Type A	*Non-TEIP*			*TEIP*		
	N	Mean	SD	N	Mean	SD
Percentage of passing scores in 9th grade exams	32 475	0.53	0.35	4 852	0.49	0.35
Mathematics 9th grade exam score	32 773	41.10	21.61	4 934	37.96	20.70
Portuguese 9th grade exam score	32 696	54.82	14.64	4 877	53.19	14.61
Percentage of passing scores in 6th grade exams	53 086	0.70	0.37	11 034	0.65	0.38
Mathematics 6th grade exam score	53 826	2.67	0.81	11 205	2.56	0.80
Portuguese 6th grade exam score	54 063	2.97	0.67	11 331	2.88	0.68
Percentage of passing scores in 4th grade exams	32 809	0.81	0.33	6 914	0.75	0.37
Mathematics 4th grade exam score	33 111	3.16	0.91	6 994	2.97	0.90
Portuguese 4th grade exam score	33 134	3.11	0.74	7 017	2.97	0.76
Successful progress in lower secondary education	34 336	0.24	0.43	5 783	0.20	0.40
7th and 8th grade without retentions	34 336	0.60	0.49	5 783	0.57	0.50
5th and 6th grade without retentions	33 912	0.58	0.49	7 329	0.57	0.50
Mother has concluded secondary education	260 037	0.11	0.32	50 653	0.10	0.30

Table A3. Did the program provide additional resources to schools?

Dep. Var.:	**Student-to-Teacher**	**Student-to-Staff**
	(1)	**(2)**
Panel A: Baseline		
TEIP schools * Post-Treatment	-0.789***	-0.564***
	(0.14)	(0.12)
Dep. Var. pre-treatment mean for treated units	9.09	6.91
ATT / Pre-treatment mean	-8.7%	-8.2%
Pre-trend test	[0.99]	[0.80]
No. of clusters	6 557	6 557
N	11 411 336	11 435 947
Panel B: By *TEIP* phase		
TEIP 2/3 schools * Post-Treatment	-0.978***	-0.755***
	(0.16)	(0.14)
TEIP 4 schools * Post-Treatment	-0.252	-0.014
	(0.32)	(0.19)
Panel C: By educational stage		
Lower Secondary: *TEIP schools* * Post-Treatment	-0.739***	-
	(0.16)	
2nd cycle: *TEIP schools* * Post-Treatment	-0.976***	-
	(0.24)	
1st cycle: *TEIP schools* * Post-Treatment	-0.406	-
	(0.31)	

Notes: Our variable of interest, *TEIP schools* * Post-Treatment, is a dummy variable with a value of 1 for students attending a *TEIP* school after it adheres to the program. Dependent Variables in columns (1) and (2) are computed at school level. The pre-trend test value in square brackets corresponds to the p-value of the test of the null hypothesis that all pre-treatment ATTs, by group and year, are equal to zero. Our analysis includes the period between the 2007/2008 and the 2017/2018 school years. Clustered standard errors at school level are presented in parentheses. Significance level at which the null hypothesis is rejected: *** 1%, ** 5%, * 10%.

Table A4. Student achievement – Exam scores using Sun and Abraham (2021)

Sub-sample:	**Mother has not concluded secondary school**		**Receiving social support – Type A**	
Dep. Var.:	**Mathematics exam score**	**Portuguese exam score**	**Mathematics exam score**	**Portuguese exam score**
	(1)	**(2)**	**(3)**	**(4)**
Panel A: 9th grade (0-100 scale)				
TEIP schools * Post-Treatment	-0.861	-0.536	-0.505	-1.474**
	(0.77)	(0.54)	(0.86)	(0.59)
N of clusters	1 182	1 182	1 157	1 157
N	485 683	484 975	162 226	161 465
Panel B: 6th grade (1-5 scale)				
TEIP schools * Post-Treatment	-0.020	0.005	0.003	0.018
	(0.02)	(0.02)	(0.03)	(0.02)
N of clusters	886	886	879	879
N	473 171	473 850	199 933	200 257
Panel C: 4th grade (1-5 scale)				
TEIP schools * Post-Treatment	0.052	0.001	0.016	-0.012
	(0.03)	(0.02)	(0.04)	(0.03)
N of clusters	914	914	889	889
N	427 781	428 513	139 546	139 785

Notes: We use the Sun and Abraham (2021) method in this table. Our variable of interest, *TEIP schools* * Post-Treatment, is an aggregation of the event-study estimates presented by the method. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Clustered standard errors at school cluster level are presented in parentheses. Significance level at which the null hypothesis is rejected: *** 1%, ** 5%, * 10%.

Table A5. Student achievement – Exam scores using the traditional TWFE

Sub-sample:	**Mother has not concluded secondary school**		**Receiving social support – Type A**	
Dep. Var.:	**Mathematics exam score**	**Portuguese exam score**	**Mathematics exam score**	**Portuguese exam score**
	(1)	**(2)**	**(3)**	**(4)**
Panel A: 9^{th} grade (0-100 scale)				
TEIP schools * Post-Treatment	-0.258	-0.742*	-0.403	-1.211***
	(0.65)	(0.39)	(0.68)	(0.40)
N of clusters	1 182	1 182	1 157	1 157
N	485 683	484 975	162 226	161 465
Panel B: 6^{th} grade (1-5 scale)				
TEIP schools * Post-Treatment	-0.008	-0.011	-0.012	-0.017
	(0.02)	(0.01)	(0.02)	(0.02)
N of clusters	886	886	879	879
N	473 171	473 850	199 933	200 257
Panel C: 4^{th} grade (1-5 scale)				
TEIP schools * Post-Treatment	0.031	-0.014	0.036	-0.004
	(0.02)	(0.02)	(0.03)	(0.02)
N of clusters	914	914	889	889
N	427 781	428 513	139 546	139 785

Notes: We use the traditional TWFE method in this table. Our variable of interest, *TEIP schools* * Post-Treatment, is a dummy variable with a value of 1 for students attending a *TEIP* school after it has adhered to the program. Our analysis includes the period between the 2006/2007 and the 2017/2018 school years. Clustered standard errors at school cluster level are presented in parentheses. Significance level at which the null hypothesis is rejected: *** 1%, ** 5%, * 10%.

Nova School of Business & Economics
Campus de Carcavelos
Rua da Holanda 1
2775-405 Carcavelos | Portugal
novasbe.pt