

# **Does attending a Selective Secondary School Improve Student Performance? Evidence from the Bicentenario Schools in Chile<sup>1</sup>.**

Pablo Araya<sup>a\*</sup> and Francisca Dusaillant<sup>b</sup>

<sup>a</sup> *Ministry of Education, Santiago, Chile;* <sup>b</sup> *Ministry of Finance, Santiago, Chile.*

**This is a preliminary draft. Please do not cite or distribute  
without permission of the authors.**

## *Abstract*

We investigate the effect of attending selective secondary schools belonging to the *Bicentenario* schools program in Chile in the period from 2011 to 2014. *Bicentenario* schools are a free education option that is intended to give vulnerable students everywhere in Chile a unique educational opportunity. By using propensity score methods such as inverse probability weighting (IPW) and inverse probability weighting combined with regression adjustment (IPWRA), we estimate the average treatment effect on the treated (ATET) using the scores from the standard 8th grade language and math tests (for schools that admit students in 7<sup>th</sup> grade). We exploit the propensity score assumptions using a rich set of covariates, including families' background and student 4th grade test scores. We find that attending a Bicentenario school improves students' performance by a range of 0.35-0.23 SD in language and 0.5-0.35 SD in math. Also, we show that a proportion of this effect is due to the outstanding performance of the new Bicentenario schools.

---

<sup>1</sup> We thank Mauricio Bravo and Juan Pablo Couyoumdjian for their participation in the UDD Bicentenario interfaculty research project, and their comments and revisions of several drafts of this work. We also thank the Chilean Ministry of Education, for facilitating us the SIMCE data. This research is the sole responsibility of the research team and should not be interpreted as representative the Ministry of Education's view.

We examine the robustness of the previous results through three different strategies. First, we implement a falsification test that shows that unobservable factors are not likely to affect the treatment or the potential outcomes. Second, we perform the estimations using the Changes-in-Changes (CIC) method, with results that are consistent with the previous estimations and that also show that lower performance quantiles obtain higher benefits. Third, we use the traditional selective public schools (*Emblemáticos*) as a more demanding control group. Results indicate that *Bicentenario* schools outperform these traditional selective public schools.

Overall, this research shows a positive effect of the *Bicentenario* schools on students' test scores, mostly driven by new *Bicentenario* schools. This paper adds new evidence to analyze the effect of selective schools in developing countries like Chile. From a policy perspective, the results suggest a careful implementation of reforms that are intended to ban academic selection in a context where the quality of public education is low.

## **Introduction**

International achievement tests indicate that Chile has the best educational system among Latin American countries (OECD, 2012; UNESCO, 2015), but it is still far from the average of developed nations. The Chilean system is characterized by huge outcome differences between government subsidized (private and public) and non-subsidized private schools (paid). Private non-subsidized schools account for 7% of the national enrolment, but their students are disproportionately represented later in the selective universities.

In 2014, the government launched a major reform to the educational system. This comprehensive education reform affected schools that currently selected their students on an academic basis, by limiting their ability to choose their pool of students. An example of these schools are the Emblemáticos (Emblematic) schools, a small group of public secondary schools (about 20) which have historically yielded academic results similar to those of private non-subsidized schools. Access to Emblematic schools has historically depended on achievement in aptitude tests and on previous school grades. Emblematic schools are located mainly in the capital and have a good reputation because a significant number of their students are admitted to the traditional Chilean universities, and some of their alumni have gained access to the political, cultural and economic elite of the country. In fact, from a historical perspective, Emblematic schools have been perceived as one of the most important means of social mobility for the lower-middle classes, and as a mechanism that helps renovate and diversify the elite (Fontaine & Urzúa, 2014). Emblematic schools represent the ideal of a meritocratic public education, one of the very few means of social mobility in the country, since most public schools do not teach the minimum contents evaluated by the university entrance test (Koljatic & Silva, 2010), leading to a low share of public school students being admitted to those selective universities. In

2015, only 18% of the students of the two most selective universities in Chile came from public schools. Three out of four of these had studied at a public Emblematic school (Eyzaguirre, 2016).

In 2010, before the recent wave of anti-selection reforms, a new program whose purpose was to replicate the Emblematic model in 60 schools across the country was launched. It was called the Bicentenario (*Bicentennial*) Program. The program responded to a perceived lack of access to adequate learning that was severely limiting the opportunities of high-achieving students. A few years after the implementation of the program, national test scores showed that Bicentennial schools were outperforming public and private voucher schools and, what is most remarkable, obtaining almost identical results as private paid schools. This effect seems to remain, even after the reform that limited student selection.

This paper estimates the impact of attending schools that joined the Bicentennial school program by studying the 8th grade test scores for three years: 2013, 2014, and 2015. Our main methodological challenge is to overcome several sources of selection bias, such as non-random assignment of students. We use propensity score methods—specifically, inverse probability weighting and regression adjustment—to estimate the treatment effects of attending a Bicentennial school and test whether our results are robust by implementing a falsification test using 4th grade test scores as outcomes and by estimating as an alternative model Athey & Imbens' (2006) changes-in-changes (CIC). Finally, we repeat the estimations with a different and more restricted control group, the Emblematics. This exercise is valuable because it is reasonable to think that unobservables, such as student motivation and parents' commitment, are similar across these educational projects, they are both performance oriented and share similar educational purposes. Also, Emblematic schools have historically been highly selective

(more than Bicentennial schools), so we can be sure that any advantage of the latter over the former is not the result of peer effects.

Our results point to large effects from attending a Bicentennial school. Propensity score estimates range from 0.33 to 0.6 SD in math, and from 0.22 to 0.43 SD in language. These magnitudes are higher in 2013 than in 2015. The falsification test confirms the credibility of these results. Additionally, our CIC estimates are similar to those obtained through the propensity score methods. They further our understanding of the phenomenon under study by showing that the lowest achievers (among the students attending the Bicentennial schools) are the ones who reap the greatest rewards from attending a Bicentennial school. Finally, we also find large effects once we restrict our control to the Emblematic schools. This would suggest that Bicentennial school performance should be explained by things different than peer composition.

Since evidence suggests that Emblematicos' results are not fully attributable to peer composition, an in-depth understanding of the causes behind its success may be of interest, in order to replicate the model, or to scale it.

The next sections are organized as follows. First, we describe the main features of the Bicentennial school program. Then we provide details on the data and methodology used in this study. We show our main results in two steps: 1) We provide propensity-score estimates as a reference point, and 2) we provide the results of different strategies to check the robustness of our estimates. Concluding remarks and policy implication are offered at the end.

## **Literature Review**

The literature on academic projects that include selection is controversial because of the problem of tracking. One of the cases most studied in developed countries is the banning of grammar schools in the United Kingdom during the sixties and the

promotion of a comprehensive school system. By using matching and instrumental variable methods, Galindo-Rueda & Vignoles (2005) find suggestive evidence that more capable students benefit from a selective system, so a comprehensive system would affect their performance negatively. Manning & Pischke (2006) suggest that the results of this reform should be interpreted carefully because banning grammar schools was not a random process across the country, and it still could involve selection. In fact, they confirmed this problem, showing that in some areas results do not differ pre-and post-treatment, even though the degree of selectivity in the area does. Further research has used more robust and proper methods or has limited the analysis to certain areas. For instance, Clark (2010) uses regression discontinuity to assess the effect of the selective system, finding positive and significant short and long-run effects. Conversely, Guyon, Maurin, & McNally (2012) analyse the effect of the comprehensive reform in Northern Ireland. They find that widening the access to grammar schools causes a net positive effect on passing national exams. Finally, for the case of the United States, Abdulkadiroğlu, Angrist, & Pathak (2014) in their study of the most selective high schools in Boston and New York, find no significant effect on admitted applicants near the cut-off of the admission process (marginal students), result that also would imply that peer effects are not strong.

Evidence from developing countries suggests that selective secondary schools may have a positive effect on test scores and college enrolment. Pop-Eleches & Urquiola (2013) use more than 2,000 regression discontinuities of Rumanian schools to find that the marginal students admitted increase their test scores by 0.1 SD, although those students struggle in adapting to a more competitive regime and receive less parental support. Jackson (2010) finds that students who attend selective schools in Trinidad and Tobago are more likely to pass national exams and, moreover, the school

effect would not be driven significantly by peer composition (Jackson, 2013). Similarly, Dustan (2010) finds that a selective group of schools in Mexico improves the performance of students near the admission cut-off by 0.19 SD.

In the case of Chile, there is still little evidence. Using propensity score methods, Allende (2015) assesses the effect of attending one of the 17 Emblematic schools on 10th grade test scores to estimate the average treatment effect on the treated (ATET). The study finds that attending an Emblematic school is associated with better test scores, and the magnitude of that effect ranges from 0.2 to 0.3 SD. However, the author suggests that this effect might only be reflecting positive peer effects.

Using similar methods, Henríquez et al. (2012) estimate the average treatment effect (ATE) of attending a private voucher school belonging to an educational network which has proven to be effective at teaching at-risk students. The study's results show that the effect (ATE) of attending a school in the network in 4th grade is higher than 0.5 SD when compared with public schools, and no lower than 0.25 SD when compared with other private voucher schools. However, the authors argue that unobserved factors, like parent motivation, given the good reputation of these schools, may be a source of overestimation of the effect.

Bucarey et al. (2014) identify the effect of attending the most traditional and selective Emblematic school in Chile. This study finds an effect of 0.25 SD on the Chilean higher education admissions test, which may potentially make a difference in university applications. Interestingly, these results suggest that the second-best option for students not making the cut-off in that school's selection process (probably another Emblematic school) is not adding value to students; or that peer effects are strong there. Using the same method, Manríquez (2016) does not find an effect of attending a girls-only Emblematic school in Chile.

Finally, Alcalde (2014) evaluates the effectiveness of the Bicentennial schools based on a model of student effort choice (scholastic effort), concluding that the restructured or existing Bicentennial schools increase their mean productivity once they enter the program, but the effect is heterogeneous. Also, she notes that the new Bicentennial schools display outstanding productivity rates; three of them are even among the 20 most productive schools of the country.

In conclusion, selective secondary schools in developing countries seem to be more effective than in developed countries. One possible explanation is the huge quality breach between the best schools (that use selection to manage their excess demand) and most schools in the system.

### **The Bicentennial School Program**

The *Bicentennial* school program is aimed at promoting the equal-opportunity principle, targeting talented students from disadvantaged families by giving them the chance to enroll in a school that increases their likelihood of access to selective higher education. The program is nationwide, in all 15 regions and in 59 out of 350 municipalities (*comunas*)<sup>2</sup> across the country<sup>3</sup>. The schools joined the program in two stages, 30 in 2011 and 30 in 2012. The program's intervention focused on the 7th grade in 2011, and on the 7th and 8th grades in 2012 and 2013. From 2014 onward, the schools no longer received pedagogical support due to the change in government administration.

---

2 A "comuna" is the smallest administrative division in the country and its local authority is the mayor. The mayor is responsible for public education. A "region" is the largest administrative division.

3 Unlike Emblematic schools located in Santiago, the nation's capital, the programme was widespread across the country.

We can classify Bicentennial schools as New, Restructured or Expandend, depending on their origins. New schools (20 in total) correspond to completely new academically-oriented schools that were opened to the public. Restructured schools (38) existed before the program was launched but implemented a change in their educational project to meet the targets of the program. Expanded schools (2) were schools that already met the targets of the program so they were only asked to join and to increase their capacity (these were not included in our analyses).

Table 1 describes the main features of the *Bicentennial* program.

### ***Implementation facts***

The Ministry of Education (MINEDUC) formed a team to monitor, support and advise the schools. At the beginning of the academic year, all schools made an initial assessment of 7th graders' math and language to obtain diagnostic information. The diagnostic test showed that the students knew barely 50% of the contents for 4<sup>th</sup> and 5<sup>th</sup> grade, so the schools added a new remedial unit to the curriculum for that year (Ministerio de Educación de Chile, 2011). If the class average did not exceed 80%, school principals were advised to repeat the test until the target was reached. Additionally, the team suggested a change in the order of the math curriculum for 7<sup>th</sup> grade; the idea was to start with easier topics, such as integers, and not fractions, facilitating the math learning process and taking into account the results of the diagnostic test. The team provided math and language teaching and learning scripts and principal and teacher training in successful strategies to improve results. A similar strategy can be found in a project implemented in the municipality of Calama, Chile (Schiefelbein, Leiva, & Schiefelbein, 2004).

Contrary to expectations, Bicentennial schools were not oversubscribed. Factors such as geographical location, city population and negative parental view of public education could explain this problem (Carrasco et al., 2014, pp. 64-65) However, admission rules were enough to attract high-achieving students. Most of the Bicentennial schools required a minimum (approximate) grade point average of 6.0 in 5th and 6th grades to apply, plus admission tests. Imposing requirements for selection was later deemed illegal, through the 2015 education reform. But the cohorts under study were admitted to the programs under the previous legal framework, that allowed selection in grades 7<sup>th</sup> and up.

## **Methods & Data**

The main challenge in our analysis is to estimate the effect of attending a Bicentennial school, taking into account the non-random assignment of students into the treatment. In fact, selection is reinforced by the program's focus on high-achieving students. Even though many Bicentennial schools were not oversubscribed (and therefore they did not have the chance to explicitly select their students), self-selection based on students' motivation or parental valuation of education, both unobservables, may be a factor given that applying to a Bicentennial school is voluntary.

## ***Data***

We use data from the SIMCE standardized national test. This dataset includes math and language test scores and a parent questionnaire that is administered in each test round. The questionnaire provides information about families' socioeconomic characteristics and other relevant information, such as the perception that the parents have of their children and of the school they attend. The data we use is a panel; we have information

on students in 4th grade (the “base” year) and four years later (8th grade), one year after some of these students started attending a Bicentennial school (we only evaluate Bicentennial Schools that start at 7th grade). The dataset is restricted to all those observations for which there is complete socioeconomic information at least in the “base” year to allow for a pre-treatment control, as will be explained later. In our analyses, we only consider students that have valid test scores both in 4th and in 8th grade, in at least one discipline (math or language), and whose parents responded to the entire parental survey in 4th grade. We also match this dataset to the final grade point average earned by the students in 5th and 6th grades, which was made available by the Ministry of Education for research purposes (both SIMCE and GPA data are nameless, but the Ministry uses a masked student identifier that is constant across its databases, allowing us to match GPAs and SIMCE scores student by student).

### ***Propensity score methods***

The propensity score is defined as the probability of being treated based on a set of observable characteristics. It relies on the creation of a control group (ex-post) from a pool of individuals that could have received the treatment based on their observable characteristics. If this control group properly emulates the counterfactual, it would be possible to estimate how Bicentennial students would have performed if there were no program.

We define our control group following the procedure of Allende (2015) in his study on *Emblematic* schools. Given that 47 out of 60 *Bicentennial* schools receive their students in 7<sup>th</sup> grade, these students belonged to a different school in 4th grade. Our data allow us to retrace them and identify their 4<sup>th</sup> grade classmates, so we construct our control group from these former classmates. 10 *Bicentennial* schools that do provide instruction for lower grades were excluded from our analysis. The control group is

further restricted to students who had a final grade point average above 5.0 in 5<sup>th</sup> and 6<sup>th</sup> grade and are categorized, according to the SIMCE test, as “elementary”<sup>4</sup> in math or language in 4<sup>th</sup> grade. Therefore, the counterfactual comes from individuals in the control group who meet the requirements to apply to a *Bicentennial* school and who are similar to their treated counterparts in terms of several relevant observables described below.

Propensity score methods rely on strong assumptions. The assumption of conditional independence or *unconfoundedness* entails no unobserved variables associated with either the potential outcomes or the treatment (Imbens & Wooldridge, 2009, p. 26). It implies that, after controlling for the set of covariates, the assignment to the treatment is as good as random. Second, there is the assumption of overlap or common support. This assumption says that it is possible to find individuals with the same propensity score both in the treatment and in the control group.

The most popular propensity score estimation method is propensity score matching (PSM), roughly consisting of comparing the average outcome of individuals that were and were not treated, but had similar propensity scores (Heinrich, Maffioli, & Vázquez, 2010). Despite its popularity, this strategy has some limitations (Khandker, B. Koolwal, & Samad, 2009). We decided, therefore, to use two additional propensity score estimation methods that address them (Lara et al., 2011). Inverse Probability Weighting (IPW) is a weighted regression model that uses the inverse of the propensity scores (PS) as weights that, theoretically, counteract the effect of selection (Imbens & Wooldridge, 2009b). Inverse-probability-weighted regression adjustment (Double-Robust) is an improved version of IPW (Cerulli 2015, p.133).

---

<sup>4</sup> SIMCE scores are divided into three categories: Insufficient, Elementary, and Appropriate.

For all our models, Propensity scores are constructed using a set of variables measured before the beginning of the program. Some of these were measured at the individual level and others at the school level. Among the former are gender, the student's grade point average in 5<sup>th</sup> and 6<sup>th</sup> grades, the student's language and math SIMCE scores in 4<sup>th</sup> grade, the mother's education level (in years), the father's education level (in years), monthly household income (in thousands of Chilean pesos<sup>5</sup>), a dummy for whether the student attended pre-school, a dummy variable indicating whether the student repeated a grade, a dummy variable for whether the family expected the student to reach higher education, and the number of books at home. At the school level, we include the type of school the student attended (private, private voucher or public) in 6<sup>th</sup> grade, a dummy for indicating whether the school is located in a rural zone and total school enrolment. We included an interaction between the average of 5<sup>th</sup> and 6<sup>th</sup> grades and type of school to capture previous schools' quality and academic demand differences. We estimate the propensity score<sup>6</sup> using a logit model and impose the common support restriction following the recommendations of Garrido et al. (2014). We also assessed the quality of our control group. For instance, the degree of overlap between treated and untreated was assessed visually (figure available from authors upon request), and the imbalance was checked across covariates following the recommendations of Garrido et al. (2014, p. 1708). We considered that standardized differences no higher than one-fourth of a standard difference to be an acceptable measure of group balance (Imbens & Wooldridge, 2009a).

---

<sup>5</sup>US\$ 1 = 500 Chilean pesos in 2010.

<sup>6</sup> We implement the propensity score using the PSCORE command (Becker & Ichino 2002). The advantage of using PSCORE is the balanced covariate obtained across treatment and non-treatment groups and within strata and the resulting optimal balance (Garrido et al., 2014).

We provide estimates of the average treatment effect on the treated (ATET) by type of *Bicentennial* school (new and restructured) to see if there are different effects between both versions of the project.

#### *Robustness checks*

The key assumption for matching and weighting methods is unconfoundedness, namely, that selection depends only on observables. Imbens & Wooldridge (2009) suggest a falsification test to examine the reliability of this strong assumption. Falsification requires the use of a new outcome variable, similar to the outcome of interest, but that should not affect the treatment. In our case, pre-treatment test scores in 4th grade meet these requirements. If the unconfoundedness assumption is true, the treatment should not have any effect on SIMCE 4th grade test scores given that we focus on the Bicentennial schools that start at the 7th grade.

We also exploit our data and create a panel to implement a Changes-in-Changes (CIC) analysis that exploits variation across treatment groups and time periods. This model, initially proposed by Athey & Imbens (2006), is a generalization of the traditional differences-in-differences method that does not impose the common trend assumption (namely, observations in the treatment and control group would behave in the same pattern in absence of the treatment). CIC analysis is valuable because it estimates the distribution of unobservables under two assumptions: a higher level of unobservables implies higher outcomes, and unobservables remain constant over time. To implement CIC estimates, we follow Melly & Santangelo (2015), who suggest a flexible semi-parametric estimator based on quantile regressions, which allows identification of the effect along the whole distribution. For these estimations, we include the same covariates as these used when estimating through propensity score

models, and we add covariates that capture changes in peer composition, such as income and parents' level of education.

Finally, we repeat our estimations using a more demanding control group: the Emblematic schools. As we mentioned in the Introduction, the Bicentennial program is inspired by the Emblematic school model, so it is reasonable to think that unobservable variables like student motivation and parent commitment would be similar for Emblematic schools and this new control group. Moreover, since the Emblematic schools are located in the capital city (the most populated city of the country) while the Bicentennial schools are scattered throughout Chile, sometimes in rural areas, Emblematic schools are better positioned to conduct a stringent selection process. Emblematic students applied to their schools with 5th and 6th-grade averages that were almost the same than those of Bicentennial students. On the other hand, Emblematic students outscored by an average of 20 4<sup>th</sup> grade SIMCE points (almost 0.5 SD) their Bicentennial counterparts. Moreover, average families' income and years of education of parents are higher in Emblematic schools. This is important because any advantage of the Bicentennial schools over the Emblematic schools can hardly be interpreted as the former being more successful at "cream skimming" or achieving, through selection, a "better" peer mix in order to exploit positive peer effects.

## **Main Results**

### ***Probability of attending a Bicentennial School***

Table 2 shows the results of the logistic regression model to predict the probability of attending a Bicentennial school for our three different cohorts of students. To facilitate interpretation, we provide the marginal effects. Results show positive marginal effects of gender (male), particularly significant for the second cohort (2014). These results are

in line with evidence that indicates that females tend to avoid the more competitive environments like those found in elite education (Koch, Nafziger, & Nielsen, 2015). 5th and 6th grades (GPAs) are significant in explaining enrolment only for public schools (in 2013 and 2014) and voucher schools (but only in 2014). Also, contrary to our expectations, previous test scores are not good at predicting enrolment. For instance, marginal effects for 4th grade SIMCE math scores are small and significant only at the 90% level in the 2013 cohort, at the 99% level in 2014, but not significant in the third one (2015). The probability of attending a Bicentennial school decreases (slightly) with higher SIMCE 4th grade language scores. These figures reinforce the idea that the program serves a group of good students who are better in terms of school grades, but who are not necessarily superior in terms of overall performance in national test scores.

Parents' educational level is positively correlated with attending a *Bicentennial* school. The marginal effect is higher for mothers than fathers. The probability of being admitted to a *Bicentennial* school is about 4% higher for students whose parents think that their children will go to university. These results are consistent with the idea that parental expectations and educational level are associated with educational decisions. Interestingly, students from lower income families have a higher probability of getting a place in a *Bicentennial* school, which indicates that the program's focalization criteria are met: students from poorer families are the ones receiving the treatment.

At the school level, results show that coming from a rural school increases the students' probability of attending a Bicentennial school. This may be true because many rural schools only provide instruction through 6th grade. These students are forced to find an educational alternative for their last years of primary education (7th and 8th) and secondary education, and the Bicentennial schools offer them a real alternative.

Alternatively, in an urban setting, it is more common to find schools that offer instruction for the complete primary and secondary cycles.

### ***Average Treatment Effect on the Treated (ATET)***

Table 3 presents the ATET estimates on attending a *Bicentennial* school based on the SIMCE math and language 8<sup>th</sup> grade test scores for our three different cohorts. Four facts can be highlighted from the results. First, student test scores seem to improve significantly when attending a *Bicentennial* school. For language, the ATET ranges from 0.42 to 0.22 SD; and for math, from 0.59 to 0.30. From the viewpoint of Murnane & Willett (2010), most of the effects identified should be considered large since they exceed their established threshold of 0.3 SD. Second, IPW estimates are slightly higher than Double-Robust (DR) ones. The DR method captures changes in peer composition (in terms of family income and parents' educational level), and these may explain this difference. The third fact is a decreasing pattern in the magnitude of the *Bicentennial* school's effect, but with a sharper fall in 2015. We are not able to fully explain this trend, but one of the hypotheses that can be studied in future research is whether it is related to changes in the program's leadership, to the school being more motivated about the program in the initial stages, or to changes in admission policies. We also suggest that the lower results in 2015 may be influenced by the loss of support from the Ministry of Education starting in 2014 after a new government administration eliminated the pedagogical unit in charge of supporting *Bicentennial* schools. This meant that in 2014, schools did not administer the diagnostic test or the follow-up tests to check if learning was being achieved, and they also did not receive pedagogical advice. Nonetheless, it should be noted that the effects identified in 2015, while lower than those of previous years, were still positive and, at least in the case of mathematics, large.

Finally, all estimations show that the Bicentennial schools are relatively more effective in boosting math scores than language scores. We believe that the nature of teaching and learning those subjects may explain the differences. For instance, language may be harder to teach, and literacy development may be influenced by the home literacy environment in a higher degree than maths. Concretely, the Bicentennial maths program offered a clearer strategy and more consistent teaching methods by rearranging the contents of the 7th and 8th grade curricula. Similar improvements for language were not offered.

Tables 4 and 5 show estimates by type of Bicentennial school: new or restructured, respectively. The two methods support the fact that new Bicentennial schools obtain better results than restructured schools. Differences are clearer for the second and third cohorts. Nonetheless, even most estimations of ATET for restructured Bicentennial schools can be considered large.

### ***Robustness checks***

Table 6 shows the results of the falsification test. When studying the 4th-grade SIMCE test score "placebo" outcome, the estimated treatment effect is always non-significant and close to zero in absolute value, across methods and cohorts. Although other factors can still influence the treatment and the potential outcomes, the falsification test provides good evidence that the unconfoundedness assumption is likely to hold.

Figure 1 and 2 display CIC estimates. We observe that the CIC results are consistent with previous estimates: average quantile effects range from 0.27 to 0.44 SD for language, and from 0.32 to 0.53 for math. Unlike propensity score methods, CIC results are similar for the 2013 and the 2014 cohorts, but display an important decline for the 2015 cohort. It is interesting to note that the program has a greater effect on the

lower quantiles, particularly for language test scores. For instance, for Q1 we find effects of around 1 SD in two cohorts, implying effects four times greater than for quantile Q9. In contrast, math test scores do not function too differently across quantiles.

Since the CIC method takes into account the distribution of unobservables (under some assumptions) and controls for changes in peer composition in terms of parents' education, expectations and income, it is allowing us to control, to some extent, for the share of the peer effects that may be responsible for the better results at *Bicentennial* schools. Finally, we compare the performance of *Bicentennial* schools to *Emblematic* schools. *Emblematic* students applied to their schools with 5<sup>th</sup> and 6<sup>th</sup>-grade grade point averages virtually identical to those of *Bicentennial* students, but they obtained, on average, test scores 0.5 SD higher than *Bicentennial* students in 4<sup>th</sup> grade. Moreover, family income and parents' education are higher in *Emblematic* schools. Therefore, this control group helps us control for unobservables such as student and parent motivation. This, since the educational projects are similar: academic-oriented with a focus on high achievement for earn a place at the best universities.

Table 7 shows that *Bicentennial* schools perform better than *Emblematic* schools. For language, estimates fluctuate between 0.4 and 0.3 SD; for math, they range from 0.5 to 0.3 SD. In fact, these last results are similar to those obtained with the original control group.

Certainly, one concern about using *Emblematic* schools as a control group is that some of these students missed school days because they joined 2011-2013 student protests about the quality of education. Evidently, protests may negatively affect performance in the standardized tests, but it is difficult to argue that they systematically

affected performance three years in a row, especially considering that there were no long-lasting student protests after 2013.

## **Discussion.**

We have estimated the ATET effect of attending schools that participated in the Bicentennial school program. The results show great benefits for students who are admitted to a Bicentennial school. We confirm that our results are consistent using two different propensity score methods and additional robustness checks. Also, our results show a downward pattern, namely, 2013 results are higher than 2015 both in language and in math. Some factors could have influenced these results, but we suggest that the loss of pedagogical support from the Ministry of Education could have affected the performance of Bicentennial schools, mainly in the 2015 cohort. This should be investigated further through future research.

Although we are not identifying each of the mechanism explaining the results of the Bicentennial program, descriptions of the implementation of the program (Carrasco et al., 2014) may shed light on some of these. First, the educational project is academic-oriented and explicitly states as an objective of the program the preparation of their students' to undertake higher education. This would naturally affect performance in a standardized test such as SIMCE. As a matter of fact, the program offered until 2014 a systematic assessment of student learning that has been evaluated positively by school principals and teachers. Also, Bicentennial schools provide support to students who struggle with some subjects in the form of extra classes and tutorials. There is evidence that corroborates that the schools achieved desirable results over and above standardized test scores. For instance, the attendance level improved and the desertion rate decreased 5 percentage points once the program started. This may be a demonstration of improved

families' commitment; in fact, some families expressed satisfaction with the new school regime even though the schools were currently more demanding.

Another mechanism that might be explaining these results is peer effect. Duflo, Dupas, & Kremer, (2011) reinforce the idea that tracking by ability is not bad per se since students who belong to a "less disperse" class get better results than students who are randomly assigned to a class.. Nonetheless, since we find that *Bicentennial* schools perform better than the at least equally selective *Emblematic* schools, it would be difficult to sustain that the academic results of *Bicentennial* schools are driven by academic selection and peer effects only.

The large effect of attending a *Bicentennial* school is comparable to the results found for a privately run network of schools in Chile, as described in Henríquez et al. (2012). Most of the schools of the network are free of charge and do not select explicitly, but there may be some self-selection at play. Resorting to the qualitative assessment that made of these schools, it is possible to find common areas between the program and Bicentennial's. First, both are greatly focused on student learning and they seem to share similar institutional goals. Second, in both the academic standards are clearly defined. Third, student's performance is monitored by a centralized team that identifies the critical points in students' learning process and finds solutions in conjunction with the schools' principals.

This study complements previous studies and confirms the productivity differences between restructured and new Bicentennial schools found in Alcalde (2014). One plausible reason for the relative advantage of new Bicentennial schools is team selection. Principals and teachers of the new schools were hired after a competitive selection process, a fact that may have helped to create an organizational culture that supports the new and demanding academic project of the school.

From a policy perspective, the program seems to add value to a downcast Chilean public education system (Carrasco et al., 2014). Our estimates provide only a short-term effect of the program, so schools' performance should be tracked to evaluate whether the program's strategies are sustainable over time and in other settings.

## References

- Abdulkadiroğlu, A., Angrist, J., & Pathak, P. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1), 137–196. <https://doi.org/10.3982/ECTA10266>
- Alcalde, P. (2014). *Selective schools and the effect of class composition. Evaluation of the Bicentenario Schools in Chile*. Retrieved from [http://jdi.econ.queensu.ca/sites/default/files/14-06-09\\_Selective Schools Program.pdf](http://jdi.econ.queensu.ca/sites/default/files/14-06-09_Selective%20Schools%20Program.pdf)
- Allende, C. (2015). *Efectividad de los Liceos Públicos de Excelencia en Chile [Emblematic schools' effectiveness in Chile] (Master's Thesis)*. Universidad de Chile. Retrieved from [http://repositorio.uchile.cl/bitstream/handle/2250/134591/Efectividad de los liceos públicos de excelencia en Chile.pdf?sequence=1&isAllowed=y](http://repositorio.uchile.cl/bitstream/handle/2250/134591/Efectividad%20de%20los%20liceos%20p%C3%BAblicos%20de%20excelencia%20en%20Chile.pdf?sequence=1&isAllowed=y)
- Athey, S., & Imbens, G. W. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, 74(2), 431–497. <https://doi.org/10.1111/j.1468-0262.2006.00668.x>
- Bravo, D., Mukhopadhyay, S., & Todd, P. E. (2010). Effects of school reform on education and labor market performance: Evidence from Chile's universal voucher system. *Quantitative Economics*, 1(1), 47–95. <https://doi.org/10.3982/QE16>
- Bucarey, A., Úrzua, S., Jorquera, M., & Muñoz, P. (2014). The effect of attending an elite public high-school in Chile. The case of the Instituto Nacional. *Estudios Públicos*, 133(Summer), 37–68. Retrieved from [http://www.cepchile.cl/dms/archivo\\_5518\\_3530/rev133\\_ABucarey-y-otros.pdf](http://www.cepchile.cl/dms/archivo_5518_3530/rev133_ABucarey-y-otros.pdf)
- Caliendo, M., & Kopeinig, S. (2005). *Some Practical Guidance for the Implementation*

- of Propensity Score Matching* (Discussion Paper Series No. 1588). *Discussion Paper Series*. Bonn. Retrieved from <http://ftp.iza.org/dp1588.pdf>
- Carrasco, A., Gutierrez, G., Bogolasky, F., Rivero, R., & Zarhi, M. (2014). *Análisis del Estado de Implementación del Programa “Liceos Bicentenario de Excelencia.”* Santiago de Chile. Retrieved from [http://www.dipres.gob.cl/594/articles-124646\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-124646_doc_pdf.pdf)
- Cerulli, G. (2015). *Econometric Evaluation of Socio-Economic Programs*. (1, Ed.) (Vol. 49). Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-46405-2>
- Clark, D. (2010). Selective Schools and Academic Achievement. *The B.E. Journal of Economic Analysis & Policy*, 10(1). <https://doi.org/10.2202/1935-1682.1917>
- Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5), 1739–1774. <https://doi.org/10.1257/aer.101.5.1739>
- Dustan, A. (2010). *Have Elite Schools Earned their Reputation? High School Quality and Student Tracking in Mexico City*. Retrieved from <https://are.berkeley.edu/fields/ae/dustan.pdf>
- Elacqua, G. (2012). The impact of school choice and public policy on segregation: Evidence from Chile. *International Journal of Educational Development*, 32(3), 444–453. <https://doi.org/10.1016/j.ijedudev.2011.08.003>
- Eyzaguirre, S. (2016). *¿Por qué Liceos de Excelencia? [Why excellence-selective schools?]*. Retrieved from [https://www.cepchile.cl/cep/site/artic/20161219/asocfile/20161219124144/pder448\\_seyzaguirre.pdf](https://www.cepchile.cl/cep/site/artic/20161219/asocfile/20161219124144/pder448_seyzaguirre.pdf)
- Fontaine, A., & Urzúa, S. (2014). *Selección, diversidad y libertad de enseñanza [Selection, diversity and freedom of education]* (Documentos de Trabajo No. 10). Santiago de Chile: Centro Latinoamericano de Políticas Económicas y Sociales - CLAPES. Retrieved from [http://clapes.uc.cl/wp-content/uploads/2014/11/DocT-10-VFinal-Selección-diversidad-y-lib-enseñanza\\_v13\\_final\\_CLAPES\\_final-SU-y-AF.pdf](http://clapes.uc.cl/wp-content/uploads/2014/11/DocT-10-VFinal-Selección-diversidad-y-lib-enseñanza_v13_final_CLAPES_final-SU-y-AF.pdf)
- Galindo-Rueda, F., & Vignoles, A. (2005). *The heterogeneous effect of selection in*

*secondary schools: understanding the changing role of ability*. Centre for the Economics of Education, London School of Economics and Political Science. Retrieved from [http://eprints.lse.ac.uk/19440/1/The\\_Heterogeneous\\_Effect\\_of\\_Selection\\_in\\_Secondary\\_Schools\\_Understanding\\_the\\_Changing\\_Role\\_of\\_Ability.pdf](http://eprints.lse.ac.uk/19440/1/The_Heterogeneous_Effect_of_Selection_in_Secondary_Schools_Understanding_the_Changing_Role_of_Ability.pdf)

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, 49(5), 1701–1720. <https://doi.org/10.1111/1475-6773.12182>

Guyon, N., Maurin, E., & McNally, S. (2012, June 1). The effect of tracking students by ability into different schools: a natural experiment. *The Journal of Human Resources*. University of Wisconsin Press. Retrieved from <http://eprints.lse.ac.uk/45095/>

Heinrich, C., Maffioli, A., & Vázquez, G. (2010). *A Primer for Applying Propensity-Score Matching*. Retrieved from <http://publications.iadb.org/handle/11319/1681>

Henríquez, F., Lara, B., Mizala, A., & Repetto, A. (2012). Effective schools do exist: low-income children's academic performance in Chile. *Applied Economics Letters*, 19(5), 445–451. <https://doi.org/10.1080/13504851.2011.583208>

Hsieh, C.-T., & Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics*, 90(8–9), 1477–1503. <https://doi.org/10.1016/j.jpubeco.2005.11.002>

Illanes, G. (2014). *Efecto Pares: ¿Qué Sabemos Realmente?* Retrieved from [http://www.cepchile.cl/dms/archivo\\_5615\\_3582/pder377\\_Gillanes.pdf](http://www.cepchile.cl/dms/archivo_5615_3582/pder377_Gillanes.pdf)

Imbens, G. W., & Wooldridge, J. M. (2009a). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1), 5–86. <https://doi.org/10.1257/jel.47.1.5>

Imbens, G. W., & Wooldridge, J. M. (2009b). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1), 5–86. <https://doi.org/10.1257/jel.47.1.5>

Jackson, C. K. (2013). Can higher-achieving peers explain the benefits to attending

- selective schools? Evidence from Trinidad and Tobago. *Journal of Public Economics*, 108, 63–77. <https://doi.org/10.1016/j.jpubeco.2013.09.007>
- Jackson, K. C. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago\*. *The Economic Journal*, 120(549), 1399–1429. <https://doi.org/10.1111/j.1468-0297.2010.02371.x>
- Khandker, S., B. Koolwal, G., & Samad, H. (2009). *Handbook on Impact Evaluation*. The World Bank. <https://doi.org/10.1596/978-0-8213-8028-4>
- Koch, A., Nafziger, J., & Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115, 3–17. <https://doi.org/10.1016/j.jebo.2014.09.005>
- Koljatic, M., & Silva, M. (2010). Algunas reflexiones a siete años de la implementación de la PSU [Some reflections seven years after the implementation of the PSU - University entrance test]. *Estudios Públicos*, 120. Retrieved from [http://www.cepchile.cl/1\\_4743/doc/algunas\\_reflexiones\\_a\\_siete\\_anos\\_de\\_la\\_implimentacion\\_de\\_la\\_psu.html#.VxJ9eEftmzw](http://www.cepchile.cl/1_4743/doc/algunas_reflexiones_a_siete_anos_de_la_implimentacion_de_la_psu.html#.VxJ9eEftmzw)
- Lara, B., Mizala, A., & Repetto, A. (2011). The Effectiveness of Private Voucher Education: Evidence From Structural School Switches. *Educational Evaluation and Policy Analysis*, 33(2), 119–137. <https://doi.org/10.3102/0162373711402990>
- Manning, A., & Pischke, J.-S. (2006). *Comprehensive versus selective schooling in England and Wales: what do we know?* Centre for the Economics of Education, London School of Economics and Political Science. Retrieved from [http://eprints.lse.ac.uk/19417/1/Comprehensive\\_Versus\\_Selective\\_Schooling\\_in\\_England\\_and\\_Wales\\_What\\_Do\\_We\\_Know.pdf](http://eprints.lse.ac.uk/19417/1/Comprehensive_Versus_Selective_Schooling_in_England_and_Wales_What_Do_We_Know.pdf)
- Manríquez, B. (2016). *Efectos de los liceos públicos emblemáticos sobre el desempeño escolar el caso del Liceo N° 1 de Niñas Javiera Carrera [Emblematic schools' effects: The case of the only-girls High School N°1] (Master's Thesis)*. Universidad de Chile. Retrieved from <http://repositorio.uchile.cl/handle/2250/138585>
- Melly, B., & Santangelo, G. (2015). *The changes-in-changes model with covariates*. Retrieved from [http://www.sciencespo.fr/newsletter/actu\\_medias/12298/Dept\\_Seminar\\_paper\\_BMelly\\_Changes\\_in\\_Changes.pdf](http://www.sciencespo.fr/newsletter/actu_medias/12298/Dept_Seminar_paper_BMelly_Changes_in_Changes.pdf)
- Ministerio de Educación de Chile. (2011). Resumen: informe de implementación Liceos

Bicentenario de excelencia [Implementation Report of the Bicentennial schools program].

Murnane, R. J., & Willett, J. B. (2010). *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. Oxford University Press. Retrieved from [https://books.google.co.uk/books/about/Methods\\_Matter\\_Improving\\_Causal\\_Inference.html?id=lA0qSsQk\\_AgC&pgis=1](https://books.google.co.uk/books/about/Methods_Matter_Improving_Causal_Inference.html?id=lA0qSsQk_AgC&pgis=1)

Murnane, R., Waldman, M., Willett, J., Bos, M. S., & Vegas, E. (2017). *The Consequences of Educational Voucher Reform in Chile*. Cambridge, MA. <https://doi.org/10.3386/w23550>

Neilson, C. (2014). *Targeted Vouchers, Competition Among Schools, And The Academic Achievement Of Poor Students*. Retrieved from <https://economics.sas.upenn.edu/events/targeted-vouchers-competition-among-schools-and-academic-achievement-poor-students>

OECD. (2012). *PISA 2012 Results in Focus. What 15-year-olds know and what they can do with what they know*. Retrieved from <http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf>

Pop-Eleches, C., & Urquiola, M. (2013). Going to a Better School: Effects and Behavioral Responses. *American Economic Review*, 103(4), 1289–1324. <https://doi.org/10.1257/aer.103.4.1289>

Rothstein, J. (2009). Student Sorting and Bias in Value-Added Estimation: Selection on Observables and Unobservables. *Education Finance and Policy*, 4(4), 537–571. <https://doi.org/10.1162/edfp.2009.4.4.537>

Sapelli, C. (2010). Some Thoughts on the Evaluation of the Chilean Voucher System. *Journal of School Choice*, 4(2), 222–231. <https://doi.org/10.1080/15582159.2010.483924>

Sapelli, C., & Vial, B. (2002). The performance of private and public schools in the Chilean vouchers system. *Cuadernos de Economía*, 39(118), 423–454. <https://doi.org/10.4067/S0717-68212002011800006>

Sapelli, C., & Vial, B. (2005). *Private vs Public Voucher Schools in Chile: New Evidence on Efficiency and Peer Effects* (No. 289). *Documentos de Trabajo*.

Instituto de Economía. Pontificia Universidad Católica de Chile. Retrieved from <https://ideas.repec.org/p/ioe/doctr/289.html>

Schiefelbein, E., Leiva, R., & Schiefelbein, P. (2004). Closing the Math Learning Gap between Chile and Developed Countries. *Journal of International Cooperation in Education*, 7, 99–114. Retrieved from <http://home.hiroshima-u.ac.jp/cice/wp-content/uploads/2014/03/7-1-10.pdf>

UNESCO. (2015). *Informe de resultados TERCE: Logros de Aprendizajes*. Santiago de Chile. Retrieved from <http://unesdoc.unesco.org/images/0024/002435/243532S.pdf>

Valenzuela, J. P., Bellei, C., & Ríos, D. de los. (2013). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Education Policy*, 29(2), 217–241. <https://doi.org/10.1080/02680939.2013.806995>

Vial, B., & Sapelli, C. (2004). Peer Effects And Relative Performance Of Voucher Schools In Chile. In *Econometric Society 2004 Latin American Meetings*. Econometric Society. Retrieved from <https://ideas.repec.org/p/ecm/latm04/96.html>

Table 1. Bicentennial School Programme Description.

|                        |                                                                                                                                                                                                                                                                                                                                            |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Who could participate? | All subsidized schools. Public and private voucher schools could apply. However, public schools were the focus of the program. In fact, the final composition was only five private voucher schools and 55 public schools. The projects were presented by the owner of the voucher schools or by the mayor, in the case of public schools. |
| Size                   | Schools should open at least 80 places in 7 <sup>th</sup> grade.                                                                                                                                                                                                                                                                           |
| Benefits               | New schools could request funding of as much as US\$1.8 million to invest in infrastructure and                                                                                                                                                                                                                                            |

|                          |                                                                                                                                                                                                                     |
|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                          | educational technologies or equipment. Restructured and Expanded schools up to US\$0.9 million. The program did not provide monetary incentives to teachers.                                                        |
| Selection decision rules | The educational project (the more academic-oriented, the better); specific initiatives to attract disadvantaged students; geographic criteria, teachers and principal qualifications. All factors equally weighted. |
| Targets                  | Schools should be among the top 10% in the SIMCE <sup>7</sup> test results and the top 5% in the PSU test results.                                                                                                  |

Source: Authors, based on (Carrasco et al., 2014).

Table 2. Probability of attending a Bicentennial school.

| Variables                                                                  | 2013                         | 2014                          | 2015                          |
|----------------------------------------------------------------------------|------------------------------|-------------------------------|-------------------------------|
| Gender (1 if student is male, 0 if female)                                 | 0.0048086<br>(0.0040743)     | 0.0114906 **<br>(0.0037164)   | 0.0034921<br>(0.0033537)      |
| Student's grade point average in 5 <sup>th</sup> and 6 <sup>th</sup> grade | 0.0031961<br>(0.0052058)     | 0.0029087<br>(0.0040119)      | 0.0054942<br>(0.004585)       |
| SIMCE 4 <sup>th</sup> grade math scores                                    | 0.0001125<br>(0.0000619)     | 0.0001418 **<br>(0.000055)    | 0.0000748<br>(0.0000505)      |
| SIMCE 4 <sup>th</sup> grade language scores                                | -0.0004179 ***<br>(0.000066) | -0.0003035 ***<br>(0.0000636) | -0.0001848 ***<br>(0.0000525) |
| Repeated grade (1 if YES, 0 if NO)                                         | -0.0067452<br>(0.0150128)    | -0.0231436<br>(0.0144267)     | -0.0268864 *<br>(0.0116973)   |

---

7 The SIMCE, Sistema de Medición de la Calidad de la Educación (Education Quality Measurement System), is the national standardized test administered every year in 4th grade, and in either 8th grade or 10th grade, that is, in year  $t$  SIMCE is given to 4th and 8th graders and, in year  $t + 1$ , to 4th and 10th graders.

|                                                                       |                                |                                |                                |
|-----------------------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|
| Attended preschool (1 if YES, 0 if NO)                                | 0.0028678<br>(0.0050208)       | 0.0070255<br>(0.0048608)       | 0.0195509 ***<br>(0.0051064)   |
| Students' mean family income in 4 <sup>th</sup> grade                 | -0.0000353 ***<br>(0.00000835) | -0.0000203 **<br>(0.00000701)  | -0.0000239 ***<br>(0.00000609) |
| Father's education                                                    | 0.0018415 **<br>(0.000763)     | 0.0005231<br>(0.0006868)       | 0.0021609 ***<br>(0.0006092)   |
| Mother's education                                                    | 0.0034967 ***<br>(0.0008296)   | 0.0007116<br>(0.0007374)       | 0.0353301 ***<br>(0.0068213)   |
| Expectations: University (1 if YES, 0 if NO)                          | 0.0401025 **<br>(0.0144066)    | 0.0386552 *<br>(0.0153054)     | 0.0353237 ***<br>(0.0068217)   |
| Expectations: Technical education (1 if YES, 0 if NO)                 | 0.0036729<br>(0.0162196)       | 0.0124513<br>(0.0172669)       | 0.0098759<br>(0.0080935)       |
| Number of books at home (1 if books > 50, 0 if none)                  | 0.0068242<br>(0.0051686)       | 0.015277 ***<br>(0.0044984)    | 0.0053142<br>(0.0040981)       |
| Public school (1 if public, 0 if not)                                 | -0.190586 ***<br>(0.0570529)   | -0.0989666<br>(0.0520601)      | -0.0902198 *<br>(0.0460619)    |
| Private (non-voucher) school (1 if private, 0 if not)                 | -0.5587032<br>(0.3172543)      | 0.5133294 *<br>(0.2489128)     | 0.4011762<br>(0.2788605)       |
| Rural (1 if rural, 0 if urban)                                        | 0.0416621 ***<br>(0.0070617)   | 0.0179256 **<br>(0.0061652)    | -                              |
| 6 <sup>th</sup> grade school enrolment                                | -                              | -0.0000568 ***<br>(0.00000472) | -0.0000521 ***<br>(0.00000377) |
| Interaction 5 <sup>th</sup> /6 <sup>th</sup> grades x public school   | 0.0116699 *<br>(0.0052384)     | 0.0111845 **<br>(0.004048)     | 0.0072465<br>(0.0046106)       |
| Interaction 5 <sup>th</sup> /6 <sup>th</sup> grades x public vouchers | 0.0074082<br>(0.0052417)       | 0.0088484 *<br>(0.0040486)     | 0.004783<br>(0.0046067)        |
| <i>N</i>                                                              | 19,890                         | 23,255                         | 29,461                         |
| <i>Pseudo R</i> <sup>2</sup>                                          | 0.1111                         | 0.1286                         | 0.1192                         |

Standard errors in parentheses - \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Table 3. ATET estimation bicentennial school

| Method                        | ATET - Language 8 <sup>th</sup> grade |                      |                      | ATET - Math 8 <sup>th</sup> grade |                      |                      |
|-------------------------------|---------------------------------------|----------------------|----------------------|-----------------------------------|----------------------|----------------------|
|                               | 2013                                  | 2014                 | 2015                 | 2013                              | 2014                 | 2015                 |
| Inverse Probability Weighting | 0.428***<br>(0.0341)                  | 0.363***<br>(0.0343) | 0.244***<br>(0.0421) | 0.596***<br>(0.0519)              | 0.519***<br>(0.0597) | 0.368***<br>(0.0493) |
| Double-Robust                 | 0.409***<br>(0.0355)                  | 0.324***<br>(0.0367) | 0.222***<br>(0.0446) | 0.531***<br>(0.0588)              | 0.479***<br>(0.0649) | 0.335***<br>(0.0455) |

Standard errors in parentheses --- \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Table 4. ATET estimation new bicentennial school.

| Method                        | ATET - Language 8 <sup>th</sup> Grade |                      |                      | ATET - Math 8 <sup>th</sup> Grade |                      |                      |
|-------------------------------|---------------------------------------|----------------------|----------------------|-----------------------------------|----------------------|----------------------|
|                               | 2013                                  | 2014                 | 2015                 | 2013                              | 2014                 | 2015                 |
| Inverse probability weighting | 0.446***<br>(0.0683)                  | 0.422***<br>(0.0587) | 0.347***<br>(0.0447) | 0.642***<br>(0.0832)              | 0.647***<br>(0.0814) | 0.479***<br>(0.0687) |
| Double-Robust                 | 0.430***<br>(0.0658)                  | 0.382***<br>(0.0591) | 0.319***<br>(0.0455) | 0.571**<br>(0.0867)               | 0.603***<br>(0.0788) | 0.443***<br>(0.0668) |

Standard errors in parentheses --- \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Table 5. ATET estimation restructured bicentennial school.

| Method                        | ATET - Language 8 <sup>th</sup> Grade |                      |                     | ATET - Math 8 <sup>th</sup> Grade |                      |                      |
|-------------------------------|---------------------------------------|----------------------|---------------------|-----------------------------------|----------------------|----------------------|
|                               | 2013                                  | 2014                 | 2015                | 2013                              | 2014                 | 2015                 |
| Inverse probability weighting | 0.420***<br>(0.0414)                  | 0.334***<br>(0.0386) | 0.194**<br>(0.0611) | 0.573***<br>(0.0759)              | 0.453***<br>(0.0799) | 0.314***<br>(0.0554) |
| Double-Robust                 | 0.398***<br>(0.0472)                  | 0.294***<br>(0.0391) | 0.176**<br>(0.0551) | 0.509***<br>(0.0765)              | 0.414***<br>(0.0843) | 0.283***<br>(0.0673) |

Standard errors in parentheses --- \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Table 6. Falsification test.

| Method                        | ATET - Language 4 <sup>th</sup> grade |                       |                      | ATET - Math 4 <sup>th</sup> grade |                        |                      |
|-------------------------------|---------------------------------------|-----------------------|----------------------|-----------------------------------|------------------------|----------------------|
|                               | 2013                                  | 2014                  | 2015                 | 2013                              | 2014                   | 2015                 |
| Inverse probability weighting | -0.00160<br>(0.00617)                 | -0.00264<br>(0.00488) | 0.00163<br>(0.00432) | -0.00348<br>(0.00561)             | -0.000882<br>(0.00582) | 0.00333<br>(0.00497) |
| Double-Robust                 | -0.00427<br>(0.0108)                  | -0.00783<br>(0.00955) | 0.00385<br>(0.00649) | 0.00377<br>(0.00790)              | 0.00631<br>(0.00816)   | 0.00359<br>(0.00606) |

Standard errors in parentheses --- \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Table 7. ATET estimation bicentennial school (Control group: Emblematic Schools).

| Method                        | ATET - Language 8 <sup>th</sup> grade |                      |                     | ATET - Math 8 <sup>th</sup> grade |                     |                    |
|-------------------------------|---------------------------------------|----------------------|---------------------|-----------------------------------|---------------------|--------------------|
|                               | 2013                                  | 2014                 | 2015                | 2013                              | 2014                | 2015               |
| Inverse Probability Weighting | 0.352 **<br>(0.128)                   | 0.310***<br>(0.0915) | 0.282**<br>(0.0877) | 0.377 **<br>(0.134)               | 0.520**<br>(0.186)  | 0.301*<br>(0.121)  |
| Double-Robust                 | 0.405***<br>(0.0998)                  | 0.290***<br>(0.0713) | 0.319**<br>(0.0998) | 0.447***<br>(0.101)               | 0.500***<br>(0.127) | 0.332**<br>(0.112) |

Standard errors in parentheses --- \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Figure 1. CIC estimates Language

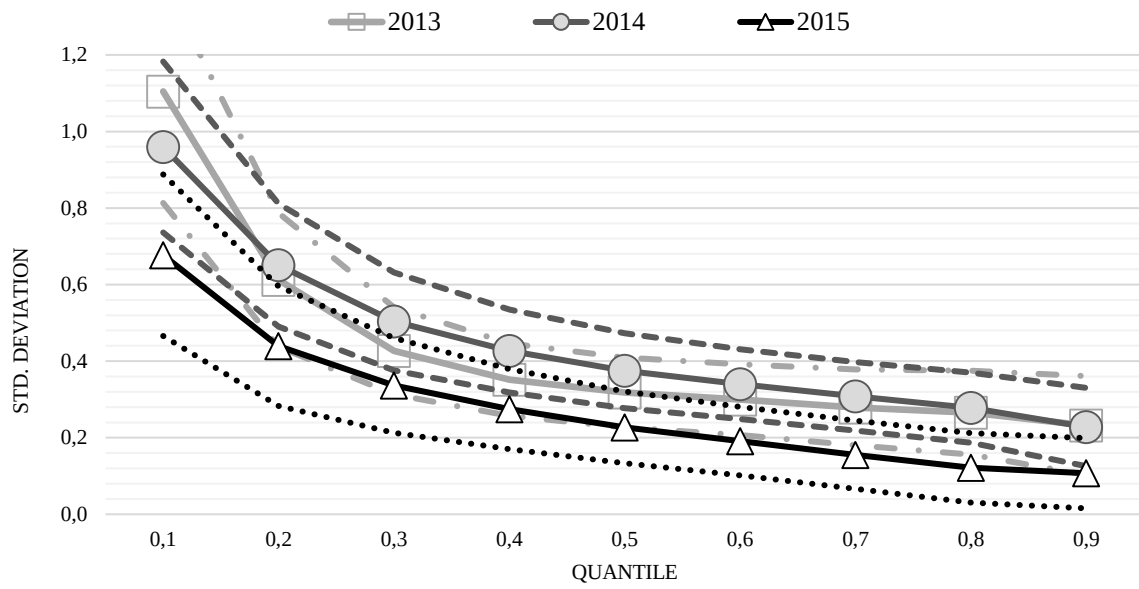


Figure 2. CIC estimates math.

