The Desgin of Vouchers and Schools' Strategic Behavior: The Case of Chile

Cristián Sánchez*

First version: May 2016
This version: October 2019

Abstract

I use rich administrative data from Chile and a difference-in-differences strategy to show that the positive effects on test scores of a recent targeted voucher reform found in the literature need to be taken with caution, as I find that the reform significantly decreased the likelihood that low-performing students take the national standardized tests. Specifically, low-performing students are 14.7 percentage points less likely to take the national exams four years after the introduction of the reform, while high-performing students are as likely to take the exams after the introduction of the reform as they were before. Such result cannot be explained by the observed increase in class attendance attributed to the program, but rather suggests a strategic response from schools to the requirement of constantly increasing performance in order to secure the receipt of the new subsidy. The ultimate consequence of this strategic behavior is that the resulting observed distribution of standardized test scores, an instrument that is used by the government to guide many of its policies, and by families to guide their enrollment decisions, may no longer be an accurate representation of the actual distribution, leading to inefficiencies in the allocation of resources.

*Centro de Investigación Económica and Department of Economics, ITAM, e-mail: cristian.sanchez@itam.mx. I am grateful for comments from Matteo Bobba, Fiona Burlig, Julián Cristiá, Clément De Chaisemartin, Ricardo Espinoza, Judy Hellerstein, Adrienne Lucas, Rachael Meager, Kensuke Teshima, Sergio Urzúa, and Tom Vogl, as well as seminar participants at ITAM, SECHI 2019, NEUDC 2019, LACEA-LAMES 2019 and Inter-American Development Bank. I thank the Ministry of Education and the Agencia de Calidad de la Educación in Chile for providing the data for this paper. All errors are my own.

1 Introduction

The educational literature that studies education markets that include voucher subsidies and school choice often focuses its attention on the demand side of these markets (i.e. students, parents).[1] They typically answer important policy questions, such as what are the gains in academic performance and achievement of attending a private school versus a public one, or what are the factors that determine parental school choice. However, in most of the cases, voucher policies involve conditions and changes in regulation that also affect the supply side of the education markets. It thus becomes crucial to understand how schools react to changes in policies if we want to have a clear picture of all the effects of such policies.[2] In this paper, I study how schools responded to a recent targeted voucher reform in Chile that considerably increased the funding per-student, but that also required a rapid increase in schools' performance to secure the receipt of the subsidy.

In 2008, the Chilean government introduced a new subsidy in the form of a targeted voucher for disadvantaged students, that supplemented the existing per-student flat voucher. Schools participating in the new program receive extra funds for every disadvantaged student that they enroll, in addition to the base voucher common to all students. The size of the new funds are considerable, representing about 60% the amount of the base voucher. A novelty of this reform is that, for the first time, the payment of the subsidy is contingent on the school improving its performance on standardized tests. This condition is closely monitored and regularly enforced by the government. The idea that motivates this condition is to make sure that participating schools exert enough effort to improve the quality of the education they provide.

Neilson (2013) and Correa et al. (2014) have recently documented large impacts of this targeted voucher reform on the performance of students and schools affected by the program. Specifically, Neilson (2013) finds that disadvantaged students increased their performance on standardized test scores by 0.2σ due to the targeted voucher reform.[3] Similarly, Correa et al. (2014) find that the program had an impact of 0.12–0.18σ on schools' performance on standardized test scores four years after the introduction of the reform. Though encouraging, this evidence is not in line with audit studies and papers that document large inefficiencies in the use of the new subsidies from the part of schools, and that only few institutions were able to effectively take advantage of the increased funds received through the program (de la Republica, 2012; Raczynski et al., 2013). Thus, it remains unclear what were the actions that participating schools took in order to

[1] See Angrist et al. (2002), Angrist et al. (2006), Hsieh and Urquiola (2006), Hastings and Weinstein (2008), Gallego and Hernando (2009), Rouse and Barrow (2009), Bravo et al. (2010), Bettinger (2011), and Carneiro et al. (2013), among others.

[2] See Bau (2014), and Dinerstein and Smith (2014) for two recent studies that effectively account for schools responses to policy changes when investigating the consequences of policies on educational outcomes.

[3] σ denotes standard deviation units.

increase their performance on standardized test scores.

I investigate a potential mechanism that schools can use to effectively and rapidly increase their performance on standardized test scores. National standardized exams in Chile are taken simultaneously by all students in the country in two specific days, usually in November. To take the tests, students have to attend school as they normally do in a regular day of class. Thus, an inexpensive mechanism that schools can use to increase their average performance is to strategically select a subgroup of high-performing students to take the exams. They can do so by asking some low-performing students to stay home during the exam days. Whether this behavior actually occurs remains an empirical question. What is clear is that schools that participate in the targeted voucher program have strong incentives to engage in such strategic behavior whenever such action is less costly than actually increasing the quality of education they provide, and effectively translating that higher quality into higher test scores. The large literature on education policies suggests that the latter is very costly.[4]

By exploiting rich administrative data on students' and schools' characteristics, performance, and test-taking rates, I use a difference-in-differences strategy to show that the targeted voucher reform significantly decreased the likelihood that low-performing students take the national standardized tests. Specifically, low-performing students are about 15 percentage points (20%) less likely to take the tests due to the program. Furthermore, the introduction of the program did not have an effect on high-performing students' likelihood of taking the tests. As a consequence, this specific strategic response of schools to the targeted voucher program introduced a bias in the representativeness of the test score distribution, where high-performing students are over-represented relative to the period before the implementation of the targeted program. Apart from being novel, this result is important for the implementation and evaluation of public policies for two main reasons. First, the results on the national standardized tests are used both by the government to allocate many of its educational policies (Cuesta et al., 2017), and by families to make their school choices (Gallego and Hernando, 2009; Cuesta et al., 2017). Therefore, a test score distribution that is not representative of the true underlying distribution may lead to inefficiencies in the allocation of resources. Second, this result may invalidate many of previous studies that use the national standardized test scores as an input in their analyses (see Neilson, 2013, and Correa et al., 2014, among others).

I also show that the targeted program had somewhat positive effects on schools' investment in educational inputs, especially those related to the hiring of new teachers, and that it reduced students' chronic absenteeism.

The paper is organized as follows. After this brief introduction, Section 2 describes the Chilean school system and the targeted voucher reform. Section 3 describes the data. Section 4 presents

[4]See Glewwe (2014), and Evans and Popova (2016a,b).

the identification strategy and shows the results of the paper. Section 5 concludes.

2 The Chilean School System and the Targeted Voucher Program

Schools in Chile can be organized within three main groups according to their management and financing scheme: public schools, private-voucher schools, and private-fee-paying schools. Both public and private-voucher schools are financed by a per-student voucher subsidy paid by the government directly to the schools. Private-fee-paying schools are financed by fees charged to parents, and serve the country's richest families. Today, 40% of students in elementary grades attend public schools, 52% attend private-voucher schools, and 8% attend private-fee-paying schools.

In 2008, the government introduced a new source of subsidy to complement the existing flat voucher in the form of a targeted voucher to disadvantaged students. On February of that year, the *Ley de Subvención Escolar Preferencial (SEP)* law that regulates this new subsidy was enacted, and was immediately put into practice for the 2008 academic year.[5] The law mandates that each school that participates in the program receives an additional subsidy per every disadvantaged student that they enroll. In addition, they also receive a per-disadvantaged student subsidy that depends on the share of disadvantaged students enrolled in the school, called *Subvención por Concentración (SC)*. Participation in the program is voluntary on the part of schools, and only public and private-voucher schools are eligible to join. Monitoring from the government is also an important aspect of the reform. At the moment of joining the program, each school is classified into one of three categories, that determines the level of monitoring the school receives from the government. The classification is based on schools' past standardized test scores and on a poverty index for the population that is served by the school. A higher classification implies less monitoring. In addition, all schools are required to set short- and long-term learning goals (i.e. test score achievements), which are evaluated by the government at the end of the period. Failing schools are reclassified one level down, and/or temporally suspended to receive the subsidy, with an eventual permanent exit of the system.

Table 1 displays the evolution of the monthly per-student voucher subsidy corresponding to elementary grades 1st-4th, decomposed by its different categories, for the years 2005-2011. Figure 1 complements this analysis graphically. The targeted voucher represents a considerable increase from the original flat voucher, of about 50–60% the base amount. The SC subsidy, in contrast, is almost negligible, representing only about 1% the base amount. The base voucher has experienced

[5]The academic year in Chile goes from March through December.

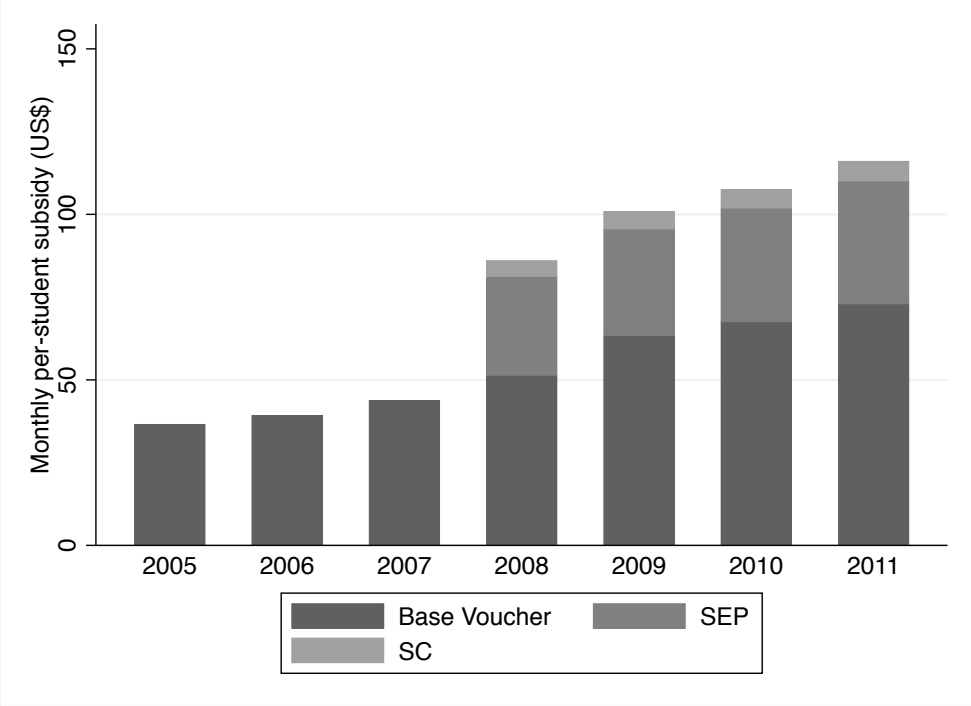
slight yearly increases during this period, with the largest increase occurred in 2009. A similar pattern is observed for the different targeted voucher categories since the introduction of the reform in 2008.

Table 1: Monthly Voucher Subsidy Decomposition for Students in 1st-4th Grades

	subsidy (US$)						
category	2005	2006	2007	2008	2009	2010	2011
base voucher	36.52	39.21	43.76	51.38	63.40	67.57	72.97
targeted voucher (SEP)	–	–	–	29.85	32.23	34.35	37.10
subsidy by concentration (SC):							
15–30%	–	–	–	2.09	2.26	2.40	2.60
30–45%	–	–	–	3.58	3.87	4.12	4.45
45–60%	–	–	–	4.78	5.16	5.50	5.94
more than 60%	–	–	–	5.37	5.80	6.18	6.68

Notes: SEP is the Spanish acronym for the targeted voucher program. SC is the Spanish acronym for the additional voucher subsidy that depends on the percentage of disadvantaged students in the school. All values are real, and are converted from Chilean pesos to US dollars according to the exchange rate of Ch$686.52 per US dollar, as of May 16, 2016. The base voucher values correspond to those for students at schools with full school shifts.

Figure 1: Size of the Vouchers for Students in 1st-4th Grades, by Category and Year



Notes: SEP is the Spanish acronym for the targeted voucher program. SC is the Spanish acronym for the additional voucher subsidy that depends on the percentage of disadvantaged students in the school. All values are real, and are converted from Chilean pesos to US dollars according to the exchange rate of Ch$686.52 per US dollar, as of May 16, 2016. The base voucher values correspond to those for students at schools with full school shifts. The SC values correspond to those for students in schools with 45–60% of disadvantaged students.

3 Data

I combine various administrative data sets of Chilean students and schools for the years 2005-2011 to form a seven-year (unbalanced) panel sample for schools, and a seven-year repeated cross-section sample for 4th grade students. I ignore private-fee-paying schools and their students, as the targeted voucher reform applies only to public and private-voucher schools. The data were obtained from the Ministry of Education and the Agencia de Calidad de la Educación, the government's agency in charge of conducting all national standardized examinations in the primary and secondary levels. The data sets include the censuses of students and schools, the census of teachers, and the annual national standardized exams for 4th graders. See appendix A for a more detailed description of each of the data sets I use.

Table 2 displays summary statistics for selected variables from this sample. Panel A shows

means and standard deviations for variables at the student level. Boys and girls are almost equally represented in the data. GPA, on a scale of 1.0–7.0, is fairly constant at 5.8 across years. Average class attendance is also constant at 93–94% across years. Both verbal and math scores increase over time, especially after the implementation of the targeted voucher reform. Test-taking rates are constant for the pre-reform period, and decrease right after that. The proportion of students that are recipients of the targeted voucher starts at 29% in 2008, increases to 45% in 2009, and stays constant at 42% thereafter. A similar pattern is observed for the proportion of disadvantaged students. Total enrollment for 4th graders slowly decreases over time, going from 253,223 in 2005 to 224,868 in 2011.

Panel B in Table 2 presents means and standard deviations for selected variables at the school level. The share of public schools decreases over the period studied, going from 63% in 2005 to 58% in 2011.[6] A similar pattern is observed for the proportion of rural schools. School inputs such class size, proportion of multigrade classes, average teacher experience, and pupil-teacher ratio show a common pattern, staying fairly constant in the pre-reform period, and decreasing thereafter. This is consistent with SEP schools spending the extra funds in performance-enhancing activities.[7] The proportion of SEP schools slightly increases over time, going from 79% in 2008 to 86% in 2011. Lastly, the total number of schools that offer 4th grade decreases from 7,963 in 2005 to 7,517 in 2011.

[6]This pattern has been shown before in other studies that analyze the Chilean school system. See, for example, Bravo et al. (2010).

[7]The decrease in average teacher experience goes in the opposite direction (whenever we believe that more years of experience implies better quality of teaching), but that could be explained by the fact that hiring new teachers necessarily implies hiring less experienced teachers if all experienced teachers are already under a contract in a school.

Table 2: Summary Statistics for 4th graders

	2005	2006	2007	2008	2009	2010	2011
A. students							
male	0.51	0.51	0.51	0.51	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
GPA	5.79	5.78	5.77	5.77	5.77	5.77	5.75
	(0.63)	(0.62)	(0.62)	(0.61)	(0.59)	(0.58)	(0.57)
avg. class attendance	0.94	0.94	0.94	0.93	0.93	0.93	0.93
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
verbal score	252.5	250.3	251.4	257.4	258.8	268.1	264.2
	(52.6)	(53.2)	(52.7)	(52.9)	(52.8)	(50.1)	(50.2)
math score	244.2	244.3	242.2	243.3	248.9	249.0	255.3
	(54.5)	(55.1)	(55.4)	(53.7)	(53.9)	(52.5)	(49.6)
test-taking rate	0.91	0.90	0.90	0.90	0.85	0.88	0.86
	(0.29)	(0.31)	(0.30)	(0.30)	(0.36)	(0.32)	(0.34)
SEP recipient	–	–	–	0.29	0.45	0.42	0.42
	–	–	–	(0.45)	(0.50)	(0.49)	(0.49)
disadvantaged student	–	–	–	0.33	0.53	0.49	0.48
	–	–	–	(0.47)	(0.50)	(0.50)	(0.50)
observations	253,223	249,344	241,006	238,196	231,074	234,353	224,868
B. schools							
public	0.63	0.62	0.61	0.60	0.60	0.59	0.58
	(0.48)	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)
rural	0.51	0.50	0.49	0.49	0.48	0.47	0.46
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
class size	24.0	23.9	23.7	23.5	23.3	22.8	22.2
	(11.3)	(11.2)	(11.2)	(11.3)	(11.2)	(11.4)	(11.8)
share of multigrade classes	0.36	0.35	0.34	0.33	0.32	0.30	0.23
	(0.45)	(0.45)	(0.45)	(0.44)	(0.44)	(0.43)	(0.40)
avg. teacher experience	18.9	19.5	19.6	18.7	18.0	18.0	15.9
	(8.2)	(8.3)	(8.5)	(8.3)	(8.2)	(8.3)	(7.9)
pupil teacher ratio	17.7	17.7	17.3	17.1	16.9	16.3	15.2
	(8.2)	(8.2)	(8.4)	(8.6)	(15.4)	(11.1)	(11.1)
SEP school	–	–	–	0.79	0.83	0.84	0.86
	–	–	–	(0.41)	(0.38)	(0.37)	(0.35)
observations	7,963	7,888	7,810	7,783	7,734	7,700	7,517

Notes: SEP is the Spanish acronym for the targeted voucher program.

4 Empirical Analysis

4.1 Identification Strategy and Preliminary Results

Despite the fact that Chile's educational system is a school choice one, where students can freely choose schools regardless of their location of residence, time and travel costs allow for the existence of differentiated local school markets. In fact, previous research has found that primary school students avoid traveling long distances to go to school, with the average student traveling less than 2.78 km (1.7 mi) (Gallego and Hernando, 2009; Chumacero et al., 2011).[8] In addition, about 90% of students attend a school that is located in the same municipality of their residence.[9] This makes municipalities a good candidate to define local school markets.[10]

The particular design of the targeted voucher reform implies that some municipalities are more affected by the program than others, depending on their share of disadvantaged students that are eligible to participate in the program. To make this point clear, take the extreme case of a municipality in which no disadvantaged student lives. The targeted voucher reform has zero effect in this municipality, in terms adding new funds, because no student is eligible to participate in the program. Conversely, a municipality in which all students come from disadvantaged families has the maximum potential of receiving additional funds.[11] Thus, it is possible to argue that different municipalities have different intensities of treatment, and that these intensities depend on the share of disadvantaged students that reside in the municipalities.

To avoid endogeneity issues when conducting my empirical analysis, I use students' municipality of residence the year before the targeted voucher program was introduced.[12] This variable is highly correlated with the current municipality of residence (88.4% of 4th graders in 2008-2011 live in the same municipality than they did in 2007), and is free of endogeneity issues because the residential decision was taken before the program was announced and implemented. Figure 2 displays the distribution of municipalities according to their share of disadvantaged students the year before the targeted voucher reform took place. Panel A presents the distribution at the

[8] Chumacero et al. (2011) calculate an average distance from home to school in the metropolitan area of Santiago of 2.57 km (1.6 mi) for 4th grade students that attend public schools, and of 2.78 km (1.7 mi) for students that attend private-voucher schools.

[9] This is the case for the sample of 4th graders that I use in my empirical analysis.

[10] See Topel (1986) and Card (2001) for other papers that use political and administrative boundaries to define local markets.

[11] I am careful to say that this increase in the funding is only "potential", because it primarily depends on the schools deciding to participate in the program. However, as I show below, the share of disadvantaged students in the municipality highly predicts the likelihood that a school chooses to participate in the targeted voucher.

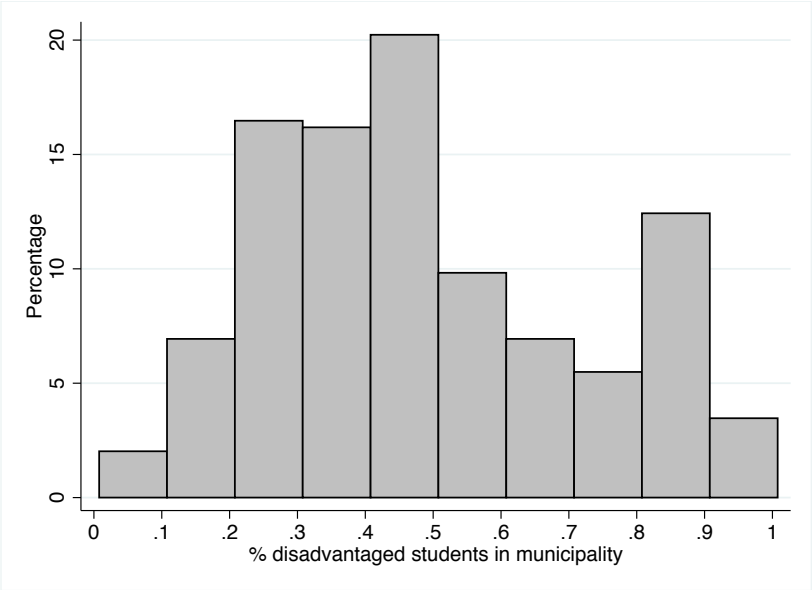
[12] As opposed to the municipality of residence before the introduction of the program, the current municipality of residence is subject to endogeneity issues via, for example, endogenous migration (Rosenzweig and Wolpin, 1988).

municipality of residence level. The support of the distribution is complete in the [0,1] range.[13] Also, about half of the municipalities have between 20% and 50% of disadvantaged students, and only few have less than 10% or more than 90% of disadvantaged students. Panel B presents the same distribution but weighted by each municipality's student population. The municipalities with the highest shares of disadvantaged students are also those with the least number of students. Also, more than half of students live in municipalities that had a share of disadvantaged students before the introduction of the reform within the range [0.1,0.4]. All in all, it is possible to argue the distribution of this "intensity of treatment" variable is well suited for my statistical analysis.

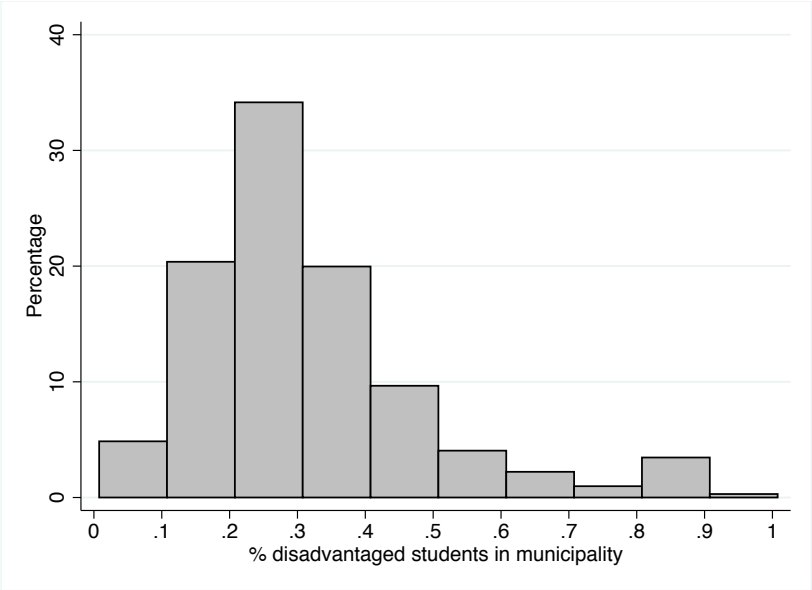
[13]The exact support of the distribution is [0.008,1].

Figure 2: Distribution of Municipalities According to their Share of Disadvantaged Students Before the Reform

A. Unweighted



B. Weighted by Municipality's Student Population

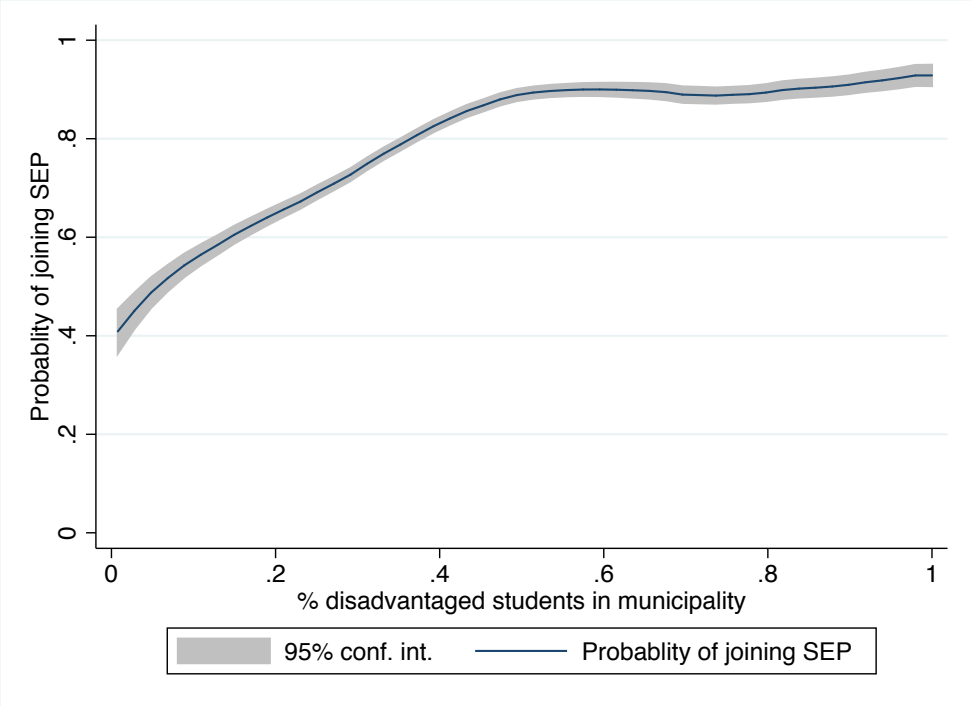


Notes: Panels A and B display histograms for the distribution of municipalities according to their share of disadvantaged students one year before the introduction of the targeted voucher reform (2007). Panel A displays the distribution at the municipality level, and Panel B presents the distribution weighted by each municipality's population of 4th grade students in the first year of the program (2008).

The basic idea of the identification strategy can be illustrated by using a standard difference-in-differences rationale. Adopting a similar strategy as in Card (1992) and Duflo (2001), I use the evidence from Figures 3 and 4 to define three categories for the share of disadvantaged students in the municipality of residence one year before the introduction of the reform.[14] Figure 3 displays a nonparametric estimation of the probability that a school joins the targeted voucher program in the first year of its implementation (2008) with respect to the municipality's share of disadvantaged students before the reform. The estimated function is monotonically increasing in the domain [0,0.5], going from 0.4 to about 0.9 in the probability, and remains fairly constant at 0.9 thereafter. Figure 4 plots a nonparametric estimation of the probability that a student attends a SEP school one year after the reform with respect to the municipality of residence's share of disadvantaged student one year before the reform. The estimated function is monotonically increasing in the domain [0,0.5], going from 0.4 to about 0.9 in the probability, remains stable in the domain (0.5,0.7], at 0.9 in the probability, and increases to almost 1 in the domain (0.7,1]. With this information at hand, I define the following three levels of intensity of treatment for local school markets (i.e. municipalities): *high*, for municipalities with more than 50% of disadvantaged students; *medium*, for municipalities with 20%-50% of disadvantaged students; and *low*, for municipalities with less than 20% of disadvantaged students. The choice of the levels responds to the fact that all students in a municipality with a high level of intensity of treatment are almost sure to attend a SEP school, and that the typical student in a municipality with a low intensity of treatment is as likely to attend a SEP school as to attend a non-SEP school. The rest of the municipalities are classified to be of medium level.

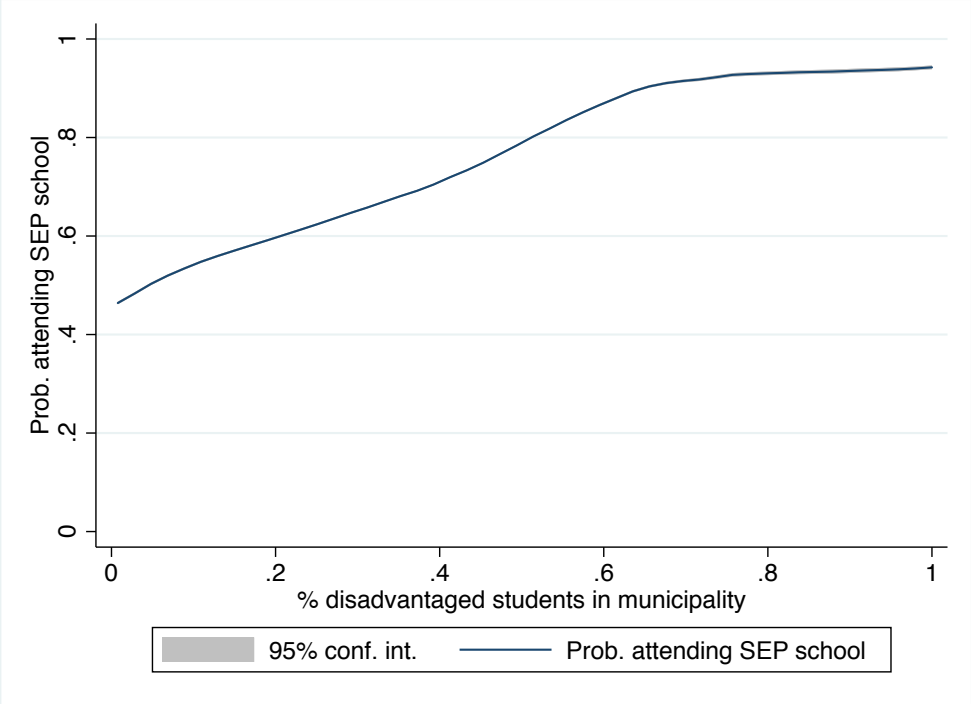
[14] Neumark and Wascher (1992) also use a similar identification strategy to estimate the effects of minimum wage laws on employment in the U.S.

Figure 3: Probability that a School Joins the Targeted Voucher Program, by Municipality's Share of Disadvantaged Students



Notes: This figure displays a nonparametric estimation of the probability that a school joins the targeted voucher reform (SEP) in the first year of its implementation (2008), with respect to the percentage of disadvantaged students in the municipality one year before the introduction of the targeted voucher reform (2007).

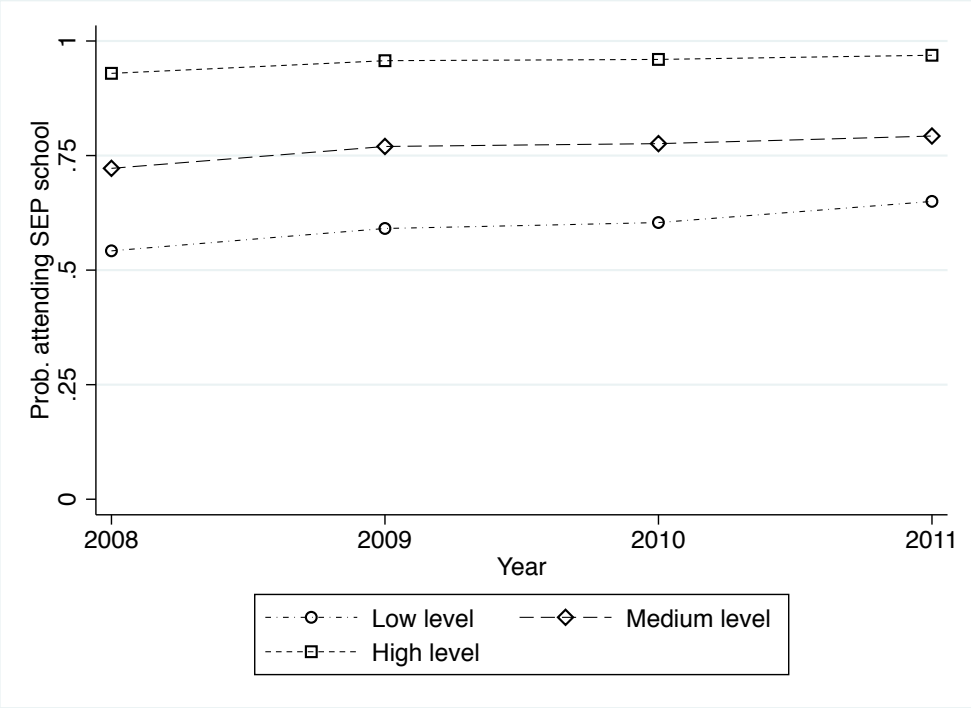
Figure 4: Probability that a Student Attends a SEP School, by Municipality's Share of Disadvantaged Students



Notes: This figure displays a nonparametric estimation of the probability that a student attends a SEP school in the first year of the targeted voucher reform (2008), with respect to the percentage of disadvantaged students in the municipality one year before the introduction of the targeted voucher reform (2007).

Figure 5 plots a dynamic version of Figure 4, where we observe that the probability that a student attends a SEP school is fairly constant over time and by intensity of treatment level, confirming the robustness of my choice for the definition of the three intensity levels.

Figure 5: Probability that a Student Attends a SEP School, by Municipality's Share of Disadvantaged Students Level Over Time



Notes: I define and use three different levels of concentration of disadvantaged students in the municipality one year before the introduction of the targeted voucher reform (2007): low, for municipalities with less than 20% of disadvantaged students; medium, for municipalities with more than 20% and less than 50% of disadvantaged students; and high, for municipalities with more than 50% of disadvantaged students. This figure displays the probability that a student attends a SEP school in the first four years of the targeted voucher reform (2008-2011), by municipality's concentration of disadvantaged students level one year before the introduction of the targeted voucher reform (2007).

Table 3 presents means and standard deviations for test scores one year before the introduction of the program (2007) and four years after the introduction of the program (2011), by intensity of treatment level. Test scores are standardized to have zero mean and standard deviation of one in 2007. Panel A presents means for verbal scores. In general, students in low-intensity municipalities have higher scores that students in medium-intensity municipalities, who in turn score higher than students in high-intensity municipalities. Average test scores increase in all three types of municipalities, but they increase more in high-intensity municipalities, followed by medium-intensity ones. Simple difference-in-difference estimators can be constructed by using the change in low-intensity municipalities as a base, and are shown in columns (4) and (5). The diff-in-diff estimate is 0.007σ (not statistically different from zero) for medium-intensity markets, and 0.091σ for high-intensity markets. In order to transform these estimates into effects attributed

to the targeted voucher reform, I divide them by the differences in the share of disadvantaged students between medium/high-intensity markets and low-intensity markets. Such shares are 15%, 31.4%, and 68.7% for low-, medium-, and high-intensity markets, respectively. Combining these results, we obtain that the average treatment effect on the treated (Abadie, 2005) of the targeted voucher reform on verbal scores is 0.04σ for students in medium-intensity markets, and 0.16σ for students in high-intensity markets. A similar analysis can be done for math scores, by using the statistics displayed in Panel B. I obtain effects attributed to the targeted voucher reform of 0.27σ and 0.31σ for medium- and high-intensity markets, respectively. These results are in line with the findings in Neilson (2013).

Table 3: Test Scores Before and After the Targeted Voucher Reform, by Intensity of Treatment Level

	intensity of treatment level			difference	
	low	medium	high	medium - low	high - low
	(1)	(2)	(3)	(4)	(5)
A. verbal					
test score in 2007	0.070	-0.010	-0.083	-0.080	-0.153
	(0.004)	(0.003)	(0.006)	(0.005)	(0.008)
test score in 2011	0.297	0.225	0.235	-0.073	-0.062
	(0.004)	(0.003)	(0.007)	(0.005)	(0.008)
difference 2011 - 2007	0.228	0.234	0.319	0.007	0.091
	(0.006)	(0.004)	(0.009)	(0.007)	(0.011)
B. math					
test score in 2007	0.115	-0.008	-0.178	-0.123	-0.293
	(0.004)	(0.003)	(0.006)	(0.005)	(0.008)
test score in 2011	0.303	0.226	0.181	-0.077	-0.122
	(0.004)	(0.003)	(0.006)	(0.005)	(0.008)
difference 2011 - 2007	0.188	0.233	0.359	0.045	0.171
	(0.006)	(0.004)	(0.009)	(0.007)	(0.011)

Notes: The sample consists in all students in 4th grade in public and private-voucher schools for the years 2007 and 2011. Test scores are standardized with respect to the year 2007. Standard errors in parentheses. I define and use three different levels of intensity of treatment: low, for municipalities with less than 20% of disadvantaged students; medium, for municipalities with more than 20% and less than 50% of disadvantaged students; and high, for municipalities with more than 50% of disadvantaged students.

I also present difference-in-differences estimates for standardized test-taking rates. Standardized tests in Chile are mandatory for schools and are scheduled to be taken simultaneously by

all 4th grade students in two specific days, usually in November, every year.[15] To take the tests, students have to normally attend class those days, just as they do in a regular day of class. Table 4 presents test-taking rates for the years 2007 and 2011, and by intensity of treatment level. Panel A displays such rates for all students. On average, test-taking rates are highest in low-intensity school markets, followed by medium-intensity markets, and then by high-intensity school markets. Also, test-taking rates decreased in all three types of municipalities between 2007 and 2011. The reduced-form diff-in-diff estimated effects of the targeted voucher reform on students' test-taking rates, using low-intensity municipalities as benchmark, are -0.7 percentage points (p.p.) for medium-intensity municipalities, and -2.5 p.p. for high-intensity municipalities. These numbers imply a treatment on the treated effect attributed to the reform of -4.3 p.p. for medium-intensity municipalities, and -4.7 p.p. for high-intensity municipalities. Panel B presents an analogous diff-in-diff exercise as in Panel A, but only for low-performing students.[16] The implied treatment on the treated effects of the targeted voucher reform on the test-taking rates of low-performing students are -12.8 p.p. for students in medium-intensity municipalities, and -10.2 p.p. for students in high-intensity municipalities. These estimates represent considerable effects of -15.7% and -13.7% in the test-taking rates of low-performing students in medium- and high-intensity municipalities, respectively. Panel C presents the same diff-in-diff exercise in test-taking rate as in panels A and B, but for high-performing students. The implied effects of the targeted voucher program are 4.9 p.p. for students in medium-intensity municipalities, and 0.7 p.p. for students in high-intensity municipalities. The combined results from panels A-C in Table 4 show that the effect of the reform on test-taking rates is entirely driven by the decrease in the test-taking rate of low-performing students.

[15] Students in 8th and 10th grades also take standardized tests, every other year.

[16] I define low-performing students as students that belong to the lowest quartile in their schools' GPA distribution. Analogously, I define high-performing students as students that belong to the highest quartile in their schools' GPA distribution.

Table 4: Test-taking Rates Before and After the Targeted Voucher Reform, by Intensity of Treatment Level

	intensity of treatment level			difference	
	low	medium	high	medium - low	high - low
	(1)	(2)	(3)	(4)	(5)
A. all students					
test-taking rate in 2007	0.911	0.903	0.883	-0.008	-0.028
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
test-taking rate in 2011	0.880	0.866	0.827	-0.015	-0.053
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
difference 2011 - 2007	-0.031	-0.037	-0.056	-0.007	-0.025
	(0.002)	(0.001)	(0.003)	(0.002)	(0.004)
B. low-performing students					
test-taking rate in 2007	0.844	0.814	0.746	-0.030	-0.098
	(0.003)	(0.002)	(0.005)	(0.004)	(0.006)
test-taking rate in 2011	0.787	0.736	0.634	-0.050	-0.153
	(0.003)	(0.002)	(0.006)	(0.004)	(0.007)
difference 2011 - 2007	-0.057	-0.078	-0.112	-0.021	-0.055
	(0.005)	(0.003)	(0.008)	(0.005)	(0.009)
C. high-performing students					
test-taking rate in 2007	0.951	0.955	0.964	0.004	0.013
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
test-taking rate in 2011	0.947	0.958	0.963	0.011	0.017
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
difference	-0.005	0.003	-0.001	0.008	0.004
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)

Notes: The sample consists of all students in 4th grade in public and private-voucher schools for the years 2007 and 2011. Standard errors in parentheses. I define and use three different levels of intensity of treatment: low, for municipalities with less than 20% of disadvantaged students; medium, for municipalities with more than 20% and less than 50% of disadvantaged students; and high, for municipalities with more than 50% of disadvantaged students. Low-performing students are students that belong to the lowest quartile in their schools' GPA distribution. High-performing students are students that belong to the highest quartile in their schools' GPA distribution.

Two important remarks can be taken from the preliminary evidence just shown. First, the targeted voucher reform increased students' academic performance (measured by test scores). Second, the reform decreased the representativeness of low-performing students in the national standardized tests. The reasons for the latter effect are unknown and hard to identifiy, but could

well relate to strategic actions taken by schools to the increased pressure from the government to rapidly increase test scores. More importantly, these results imply that the estimated effects that I and others studies find for test scores may be biased upwards.[17]

4.2 Main Results

To better exploit the variation in treatment intensity across municipalities, I generalize the strategy presented above to a regression framework. Specifically, I estimate the following equation:

$$y_{ist} = \gamma_s + \lambda_t + \delta(\text{post}_t \times \text{intensity}_s) + \varepsilon_{ist}, \quad (1)$$

where y_{ist} is the outcome of interest for student i in municipality of residence before the reform s in period t, γ_s is a fixed effect for municipality of residence before the reform, λ_t is a year fixed effect, post_t is a post-reform indicator, intensity_s is the intensity of treatment (i.e. share of disadvantaged students in the municipality of residence one year before the reform), and ε_{ist} is an error term. In the empirical implementation of this regression I cluster the standard errors at the municipality of residence before the reform level (Bertrand et al., 2004; Cameron and Miller, 2015). The coefficient of interest is δ, as it captures the effect of the targeted voucher reform. Note that the treatment parameter identified by this equation is the average treatment effect on the treated (Abadie, 2005), which in this case of a continuous treatment represents the effect on students residing in municipalities with a 100% share of disadvantaged students before the reform.

I first estimate equation (1) on test scores and test-taking rates, and check whether the results from Section 4.1 are also found in this regression framework. Panels A and B in Table 5 present the results for test scores. Column (1) presents estimates from specifications not including controls, and column (2) includes the gender of the student and the share of public schools in the current municipality of residence as covariates. The estimated effect for verbal is 0.21σ, and the effect for math is 0.35σ, both statistically significant at all conventional levels. Panel A in Table 6 presents the results for test-taking rate. Columns (1) and (2) show estimates for all students, columns (3) and (4) do so for low-performing students, and columns (5) and (6) show results for high-performing students. The estimates indicate that the targeted voucher reform decreased in 7.9 p.p. students' test-taking rate. They also show that the targeted voucher reform decreased in 14.6 p.p. low-performing students' test-taking rate. The program had no significant effect on test-taking rate of high-performing students. This evidence confirms the negative effect that the reform had on the representativeness of low-performing students in the sample of students that

[17] See Neilson (2013) and Correa et al. (2014).

take the national exams. In addition, it is no longer possible to claim that the estimated effects found for test scores in Table 5 and in other studies (Neilson, 2013; Correa et al., 2014) are free from representativeness bias.

To the question of why schools engage in such a strategic behavior, I return to the discussion from section 1. Schools that participate in the targeted program are required to meet specific academic goals in the form of average test scores at the school level. They need to comply with the academic goals in order to secure the receipt of the new subsidy. Schools are thus incentivized to increase the quality of the education they provide, and to transform that quality improvement into higher students' test scores. However, if schools find it difficult or costly to quickly raise test scores (Glewwe, 2014; Evans and Popova, 2016a,b), they may look for alternatives that may help them achieve their goals. One such alternative is to keep low-performing students to take the tests, thus automatically increasing the school's average test score without necessarily changing its quality of education. Moreover, in the current policy setting, this is an inexpensive way of increasing average test scores, as there is no penalty for schools engaging in this behavior.

Table 5: Effect of the Targeted Voucher Reform on Test Scores and Chronic Absenteeism

	(1)	(2)
A. verbal scores		
post × intensity	0.205***	0.205***
	(0.034)	(0.033)
observations	373,260	373,260
R^2	0.031	0.040
B. math scores		
post × intensity	0.352***	0.349***
	(0.038)	(0.038)
observations	373,332	373,332
R^2	0.038	0.039
C. chronic absenteeism		
post × intensity	-0.173***	-0.172***
	(0.024)	(0.024)
observations	414,299	414,299
R^2	0.029	0.030
controls	no	yes

Notes: All results come from estimation of generalized diff-in-diff regressions that use data for students one year before and four years after the introduction of the SEP reform. The intensity of treatment variable is the municipality's share of disadvantaged students. I report the estimated coefficient on the interaction between the dummy for the period after the introduction of the program and the treatment variable. Specifications with controls include the gender of the student and the share of public schools in the current municipality of residence. Standard errors are clustered at the municipality of residence before the program level. *** denotes statistically significance at 99% level.

Table 6: Effect of Targeted Voucher Reform on Test-taking Rate and Class Attendance

	students' performance group					
	all		low-performing		high-performing	
	(1)	(2)	(3)	(4)	(5)	(6)
A. test-taking rate						
post × intensity	-0.079***	-0.079***	-0.147***	-0.146***	-0.004	-0.004
	(0.013)	(0.013)	(0.023)	(0.022)	(0.012)	(0.012)
observations	414,299	414,299	108,477	108,477	99,938	99,938
R^2	0.011	0.013	0.032	0.032	0.007	0.007
B. class attendance						
post × intensity	0.028***	0.028***	0.040***	0.040***	0.019***	0.019***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
observations	414,299	414,299	108,477	108,477	99,938	99,938
R^2	0.045	0.045	0.048	0.048	0.050	0.050

Notes: All results come from estimation of generalized diff-in-diff regressions that use data for students one year before and four years after the introduction of the SEP reform. The intensity of treatment variable is the municipality's share of disadvantaged students. I report the estimated coefficient on the interaction between the dummy for the period after the introduction of the program and the treatment variable. Low-performing students are students that belong to the lowest quartile in their schools' GPA distribution. High-performing students are students that belong to the highest quartile in their schools' GPA distribution. Specifications with controls include the gender of the student and the share of public schools in the current municipality of residence. Standard errors are clustered at the municipality of residence before the program level. *** denotes statistically significance at 99% level.

Next, I investigate whether the observed effect of the targeted voucher reform on students' test-taking rate responds to a decrease in general class attendance. Panel B in Table 6 presents results from estimating equation (1) for students' average annual class attendance rate. The estimates show a positive effect attributed to the reform. The point estimates are 2.8 p.p., 4 p.p., and 1.9 p.p. for all, low-performing, and high-performing students, respectively. Additional evidence supporting this finding is presented in Panel C in Table 5, that displays the estimated effect of the targeted voucher reform on students' chronic absenteeism. Chronic absenteeism is defined as missing ten percent of a school year for any reason (Balfanz and Byrnes, 2012), and is found to be strongly linked to low academic achievement and high dropout rates.[18] The results

[18]See Balfanz and Byrnes (2012), and Gottfried (2014).

indicate that the targeted voucher reform reduced chronic absenteeism in 17.2 p.p., out of a base of 20.7% in 2007, a result that is both statistically significant and economically important. These two pieces of evidence highlight the positive effect that targeted voucher program had on attendance, an important input for academic achievement. However, such results do not help to explain the observed decrease in the test-taking rates attributed to the reform.

As a final empirical exercise, I estimate equation (1) for selected variables that measure educational inputs. I do so to investigate whether the targeted voucher reform had an effect on schools' incentive to invest in school quality. Table 7 reports the effect of the targeted voucher reform on class size, pupil-teacher ratio, % of multigrade classes, and average teacher experience in years, all measured at the school level. I weigh each observation by schools' 4th grade enrollment. Column (1) presents estimates from specifications not including controls, while column (2) includes a dummy for public school as a covariate. Class size, pupil-teacher ratio, and average teacher experience are transformed to logs, so the estimates should be interpreted as percent changes. The results suggest that schools used the extra funds they received from the targeted voucher to invest in educational inputs. Specifically, the reform reduced class size in 8.1%, reduced the pupil-teacher ratio in 5.3%, and reduced the percentage of multigrade classes in 7.8 p.p. (out of a base of 19.2% in 2007). The program also reduced the average years of experience of teachers, in about 11.6%.

Table 7: Effect of the Targeted Voucher Reform on School Inputs

	(1)	(2)
A. class size		
post × intensity	-0.083***	-0.081***
	(0.017)	(0.016)
observations	12,071	12,071
R^2	0.274	0.290
B. pupil-teacher ratio		
post × intensity	-0.054**	-0.053**
	(0.025)	(0.024)
observations	11,991	11,991
R^2	0.261	0.266
C. % multigrade classes		
post × intensity	-0.078***	-0.078***
	(0.011)	(0.011)
observations	12,071	12,071
R^2	0.171	0.171
D. avg. teacher experience		
post × intensity	-0.104*	-0.116**
	(0.055)	(0.052)
observations	11,995	11,995
R^2	0.164	0.425
controls	no	yes

Notes: All results come from estimation of generalized diff-in-diff regressions that use data for schools one year before and four years after the introduction of the SEP reform. Each school observation is weighted by school's enrollment in 4th grade. The intensity of treatment variable is the municipality's share of disadvantaged students. I report the estimated coefficient on the interaction between the dummy for the period after the introduction of the targeted voucher program and the treatment variable. Specifications with controls include a dummy for public school. Standard errors are clustered at the municipality of residence before the program level. *** denotes statistically significance at 99% level. ** denotes statistically significance at 95% level. * denotes statistically significance at 90% level.

In summary, I find that the targeted voucher program had positive effects on test scores four years after its implementation. However, those results should be taken with caution, as I also find that the reform significantly decreased the likelihood that low-performing students take the national standardized tests, a result that cannot be explained by the observed increase in class attendance due to the program. I also find that schools invested in educational inputs, namely class size, pupil-teacher ratio, and % of multigrade classes. Both the test-taking rate and the school inputs effects can certainly help explain the large effects of the reform on test scores.

5 Conclusions

In this paper, I present evidence regarding the strategic response of schools to a targeted voucher reform in Chile. Specifically, I use a difference-in-differences strategy to document that schools in local markets that were most affected by the reform engaged in the strategic selection of high-performing students to take the national standardized exams, as a way of complying with the requirement of increasing the school's average performance on standardized test scores. My results show that the reform decreased the likelihood of taking the standardized tests for low-performing students in about 14.6 percentage points four years after the introduction of the reform. Moreover, the reform did not have a significant impact on the likelihood of taking the exams for high-performing students. I also show that the reform had a positive impact on schools' investment in educational inputs, especially in those related to the hiring of new teachers.

My findings highlight the importance of taking into account the supply responses to policy changes when evaluating educational programs. The general enthusiasm generated by recent studies documenting large impacts of the Chilean targeted voucher reform on students' performance in standardized test scores must be taken with caution, as those studies do not account for schools' reactions to the reform.

References

Abadie, A. (2005): "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies*, 72, 1–19.

Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer (2002): "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment," *American Economic Review*, 92, 1535–1558.

Angrist, J., E. Bettinger, and M. Kremer (2006): "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *American Economic Review*, 96, 847–862.

Balfanz, R. and V. Byrnes (2012): "The Importance of Being In School: A Report on Absenteeism in the Nation's Public Schools," Tech. Rep. May.

Bau, N. (2014): "School Competition and Product Differentiation," .

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*.

Bettinger, E. (2011): *Educational Vouchers in International Contexts*, vol. 4, Elsevier B.V.

Bravo, D., S. Mukhopadhyay, and P. E. Todd (2010): "Effects of school reformon education and labor market performance: Evidence from Chile's universal voucher system." *Quantitative Economics*, 1, 47–95.

Cameron, A. C. and D. L. Miller (2015): "A Practitioner's Guide to Cluster- Robust Inference," *Journal of Human Resouces*, 50, 317–372.

Card, D. (1992): "Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage," *Industrial and Labor Relations Review*, 46, 22–37.

——— (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19, 22–64.

Carneiro, P., J. Das, and H. Reis (2013): "Parental Valuation of School Attributes in Developing Countries: Evidence from Pakistan," *Manuscript*.

Chumacero, R. A., D. Gómez, and R. D. Paredes (2011): "I would walk 500 miles (if it paid): Vouchers and school choice in Chile," *Economics of Education Review*, 30, 1103–1114.

Correa, J. A., F. Parro, and L. Reyes (2014): "The Effects of Vouchers on School Results: Evidence from Chile's Targeted Voucher Program," *Journal of Human Capital*, 8.

Cuesta, J. I., F. González, and C. Larroulet (2017): "Distorted Quality Signals in School Markets," *Manuscript*.

de la Republica, C. G. (2012): "Consolidado Subvención Escolar Preferencial, Informe 9/2012," .

Dinerstein, M. and T. D. Smith (2014): "Quantifying the Supply Response of Private Schools to Public Policies," .

Duflo, E. (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, September.

Evans, D. K. and A. Popova (2016a): "Cost-Effectiveness Analysis in Development: Accounting for Local Costs and Noisy Impacts," *World Development*, 77, 262–276.

——— (2016b): "What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews," *World Bank Research Observer*, 31, 242–270.

Gallego, F. A. and A. E. Hernando (2009): "School Choice in Chile: Looking at the Demand Side," .

Glewwe, P. (2014): *Education Policy in Developing Countries*, Chicago: The University of Chicago Press.

Gottfried, M. A. (2014): "Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes," *Journal of Education for Students Placed at Risk*, 19, 53–75.

Hastings, J. S. and J. M. Weinstein (2008): "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*.

Hsieh, C.-T. and M. Urquiola (2006): "The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program," *Journal of Public Economics*, 90, 1477–1503.

Neilson, C. (2013): "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," .

Neumark, D. and W. Wascher (1992): "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws," *Industrial and Labor Relations Review*, 46, 55–81.

Raczynski, D., G. Muñoz, J. Weinstein, and J. Pascual (2013): "Subvencion Escolar Preferencial (SEP) en Chile: Un Intento por Equilibrar la Macro y la Micro Politica Escolar," *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educacion*, 11.

Rosenzweig, M. R. and K. I. Wolpin (1988): "Migration, Selectivity and the Effects of Public Programs," *Journal of Public Economics*, 37, 265–289.

Rouse, C. E. and L. Barrow (2009): "School Vouchers and Student Achievement: Recent Evidence and Remaining Questions," *Annual Review of Economics*, 1, 17–42.

Topel, R. H. (1986): "Local Labor Markets," *Journal of Political Economy*, 94, S111–S143.

A Data

Below, I provide a list of each administrative and survey data set used in this paper, along with its corresponding description:

- *Registry of students, 2005-2011.*
  These data provide information on students' gender, date of birth, age, municipality of residence, type and level of education, grade, class, grade repetition status, special education status, and various characteristics of the school of attendance, such as municipality, type of administration (public, private-voucher, private-fee-paying), single/double shift schedule, and urban status.

- *Registry of students' academic performance, 2005-2011.*
  These data provide information on students' gender, date of birth, municipality of residence, type and level of education, grade, class, GPA, average class attendance, and various characteristics of the school of attendance, such as municipality, type of administration, and urban status.

- *Registry of schools, 2005-2011.*
  These data provide information on schools' municipality, type of management, urban status, address, and type and level of education offered.

- *Registry of schools' summary of enrollment, 2005-2011.*
  These data provide information on schools' municipality, type of management, urban status, male enrollment by education type and level, female enrollment by education type and level, total enrollment by education type and level, total enrollment, number of single-grade classes by education type and level, total number of single-grade classes, number of multigrade classes by education type and level, and total number of multigrade classes.

- *Registry of teachers, 2005-2011.*
  These data provide information on teachers' gender, date of birth, education degree, subject specialization, institution attended, graduation year, and duration of the degree studied. They also provide information on the characteristics of all schools in which each teacher is hired (municipality, type of management, rural status), and on the teachers' primary and secondary roles (e.g. teacher, principal, supervisor), type of contract, hours contracted, teaching hours, experience, tenure, and teaching subject and level of education.

- *Registry of schools that participate in the targeted voucher program, 2008-2011.*
  These data provide information on the characteristics of schools that participate in the targeted voucher program. Information on schools' municipality, type of management, urban status, targeted voucher classification, number of disadvantaged students that are eligible

for the targeted voucher subsidy, and number of targeted voucher beneficiary students is also available.

- *Registry of students that are eligible to participate in the targeted voucher program, 2008-2011.*
  These data provide information on the characteristics of disadvantaged students that are eligible to participate in the targeted voucher program. They provide information on students' gender, date of birth, targeted voucher participation status, level of education, grade, single/double shift schedule, and on the type of management, urban status, and targeted voucher category of the school attended by the student.

- *National standardized exams (SIMCE) for 4th graders, student-level, 2005-2011.*
  These data provide information on students' test scores for three different subjects: verbal, mathematics, and either social sciences or natural sciences, depending on the year. They also provide information on students' gender and grade.

- *National standardized exams (SIMCE) for 4th graders, school-level, 2005-2011.*
  These data provide information on schools' average test scores for verbal, mathematics, and social/natural sciences, municipality, type of mangement, socio-economic category of the population served by the school, urban status, and number of students taking the tests.

- *4th grade SIMCE's questionnaire to parents and tutors, 2005-2011.*
  These data consist in responses to a survey that parents and tutors answer during the days when the national standardized tests are taken. The survey is voluntary, though more than 90% of parents choose to respond it every year. It provides information on students' household size, house amenities, and time use, total number of books available in the household, household's total monthly income, parents and tutors' time use, education, indigenous identification, occupation, health insurance, participation in social programs, reasons for the choice of the school, beliefs on the student's future educational attainment, satisfaction with the school, knowledge of school's average performance in standardized tests, total monthly expenses related to the student's education other than tuition, and school's admission criteria, tuition, and fees.