Value Added: Estimation and Analysis

Christopher A. Neilson*

Last Updated on March 17, 2020- [See most recent version]

A main desirable characteristic of schools in the context of school choice and competition is quality, which has several ways to be estimated. Direct and commonly used proxies are students' test scores or GPAs, but broadly known is that the student ability and other characteristics are crucial components, and not all are readily observable. School quality is a measure of how much the school increases students' achievement and knowledge, controlling for all other relevant variables.

1 Estimation of Value Added

To estimate the school value added, as a measure of quality, I use students' test scores and take a large set of students' observables, including health information at birth, demographic composition of the families, parents' employment and educational levels as well as mothers' math and language college-entrance exam scores. With these variables, I define the relationship between students' achievement y_{ijt}, the student's characteristics, and the schools' ability to increase achievement q_{jt} by the following equation:

$$y_{i,j,t} = q_{j,t} + X_{i,t}\gamma + e_{i,j,t} \tag{1}$$

The estimated value of $q_{j,t}$ is the school fixed effect. It is the component of the average test score in the school that is not explained by the individual characteristics of the students. This measure of school quality captures schools' inputs such as teacher quality, infrastructure, school environment, and any other school-specific characteristic that improves achievement, measured as the average test score. To the extent that the demographic composition of the schools' students matters for test scores, these effects will also be included in the estimated school value added.

Results of the estimation are presented in Table 1, using 4th-grade students' test scores. They are three specifications, one with a basic set of students' characteristics (mother education level, income level, and gender) -**VA1**-, and two others with an extended set of characteristics -**VA2** and **VA3**-. The difference between them is that one has school-by-year fixed effects, and the other has schools fixed effect by two periods (2005-2007 and 2010-2012). The first characteristics considered are mother's level of education, which are dichotomic variables for mother with more than high school, high school, or less than high school. I add covariates that indicate if the mother took the college entrance exam (PAA), and in which decile she performed in maths and language tests. Within sociodemographics variables, we have the student's gender, marital status of her parents, quintiles for health characteristics at birth, like weight, length, and gestation weeks; type of birth (single or double), location of birth (hospital, at home), and an index if she was her mother's firstborn. I also consider indexes for the region of the country where the student born. Table 2 show the number of observations over time for the two specifications with the extended set of covariates.

*This document is an accompaning text for *Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students*. The document has benefited greatly from the help of Claudia Allende, Isabel Jacas and Maria Elena Guerrero.

Table 1: School Quality Estimation Regression

	Avg. Test Score (VA1)		Avg. Test Score (VA2)		Avg. Test Score (VA3)	
	Coef.	StdErr	Coef.	StdErr	Coef.	StdErr
Constant	0.016***	(0.000)	-0.092***	(0.000)	-0.151***	(0.000)
Mother High School	0.226***	(0.000)	0.202***	(0.000)	0.208***	(0.000)
Mother More than High School	0.429***	(0.000)	0.276***	(0.000)	0.271***	(0.000)
Male	-0.051***	(0.000)	-0.064***	(0.000)	-0.052***	(0.000)
Parents Married	-	-	0.056***	(0.000)	0.072***	(0.000)
Single Birth	-	-	0.054***	(0.000)	0.058***	(0.000)
First Born	-	-	0.056***	(0.000)	0.074***	(0.000)
Mother Took PAA	-	-	-0.102***	(0.000)	-0.086***	(0.000)
Mother PAA Math D2	-	-	0.012***	(0.006)	0.016***	(0.009)
Mother PAA Math D3	-	-	0.029***	(0.000)	0.021***	(0.001)
Mother PAA Math D4	-	-	0.047***	(0.000)	0.042***	(0.000)
Mother PAA Math D5	-	-	0.068***	(0.000)	0.070***	(0.000)
Mother PAA Math D6	-	-	0.080***	(0.000)	0.082***	(0.000)
Mother PAA Math D7	-	-	0.094***	(0.000)	0.092***	(0.000)
Mother PAA Math D8	-	-	0.096***	(0.000)	0.102***	(0.000)
Mother PAA Math D9	-	-	0.109***	(0.000)	0.113***	(0.000)
Mother PAA Math D10	-	-	0.152***	(0.000)	0.154***	(0.000)
Mother PAA Lang D2	-	-	0.080***	(0.000)	0.081***	(0.000)
Mother PAA Lang D3	-	-	0.130***	(0.000)	0.136***	(0.000)
Mother PAA Lang D4	-	-	0.174***	(0.000)	0.179***	(0.000)
Mother PAA Lang D5	-	-	0.207***	(0.000)	0.223***	(0.000)
Mother PAA Lang D6	-	-	0.236***	(0.000)	0.247***	(0.000)
Mother PAA Lang D7	-	-	0.273***	(0.000)	0.286***	(0.000)
Mother PAA Lang D8	-	-	0.305***	(0.000)	0.318***	(0.000)
Mother PAA Lang D9	-	-	0.349***	(0.000)	0.359***	(0.000)
Mother PAA Lang D10	-	-	0.431***	(0.000)	0.436***	(0.000)
Birth Weigth D2	-	-	0.034***	(0.000)	0.039***	(0.000)
Birth Weigth D3	-	-	0.049***	(0.000)	0.054***	(0.000)
Birth Weigth D4	-	-	0.059***	(0.000)	0.064***	(0.000)
Birth Weigth D5	-	-	0.063***	(0.000)	0.068***	(0.000)
Birth Gestation D2	-	-	-0.008***	(0.000)	-0.007**	(0.035)
Birth Gestation D3	-	-	-0.026***	(0.000)	-0.023***	(0.000)
Birth Gestation D4	-	-	-0.044***	(0.000)	-0.040***	(0.000)
Birth Gestation D5	-	-	-0.057***	(0.000)	-0.054***	(0.000)
Birth Length D2	-	-	0.023***	(0.000)	0.023***	(0.000)
Birth Length D3	-	-	0.030***	(0.000)	0.033***	(0.000)
Birth Length D4	-	-	0.041***	(0.000)	0.046***	(0.000)
Birth Length D5	-	-	0.057***	(0.000)	0.061***	(0.000)
Birth Location D1	-	-	-0.032***	(0.000)	-0.040***	(0.000)
Birth Location D2	-	-	-0.099***	(0.000)	-0.112***	(0.000)
Region Birth FE			✓		✓	
School by Year FE	✓		✓			
School by Group Year FE					✓	
R^2	0.30		0.31		0.28	
N Obs	2,166,730		2,164,812		1,108,152	

Note: This table presents regression results for estimates of test scores on a large vector of individual student-level characteristics. School quality is estimated as the school and year fixed effect for column (1) and (2), and as the school and year group fixed effect for column (3) (Groups are 2005-2007 and 2010-2012). Estimates of school quality have not been presented in this table.

Table 2: School Quality Estimation Regression - Observations and Missings

Year	Obs in VA Estimation	Missings	Total Obs
2005	188,849	31,702	220,551
2006	189,802	31,989	221,791
2007	187,429	27,852	215,281
2008	184,621	26,789	211,410
2009	173,490	29,898	203,388
2010	185,204	23,856	209,060
2011	178,644	26,606	205,250
2012	178,224	27,367	205,591
2013	176,068	27,655	203,723
2014	175,984	26,426	202,410
2015	173,612	28,970	202,582
2016	172,885	32,984	205,869
2017	179,540	32,771	212,311
School by Year FE	2,344,352	374,865	2,719,217
School by Group of Years FE (05-07 and 10-12)	1,108,152	169,372	1,277,524

Note: This table presents observations by year in regression showed in Table 1 results for estimates of test scores on a large vector of individual student-level characteristics. School quality is estimated as the school and year fixed effect and has not been presented in this table.

2 Stability of Coefficients

To test the stability of coefficients γ from Equation 1 before and after the policy change, I first run the regression separately for each year prior to the policy as well as the combined regression. The results are presented in Table 3.

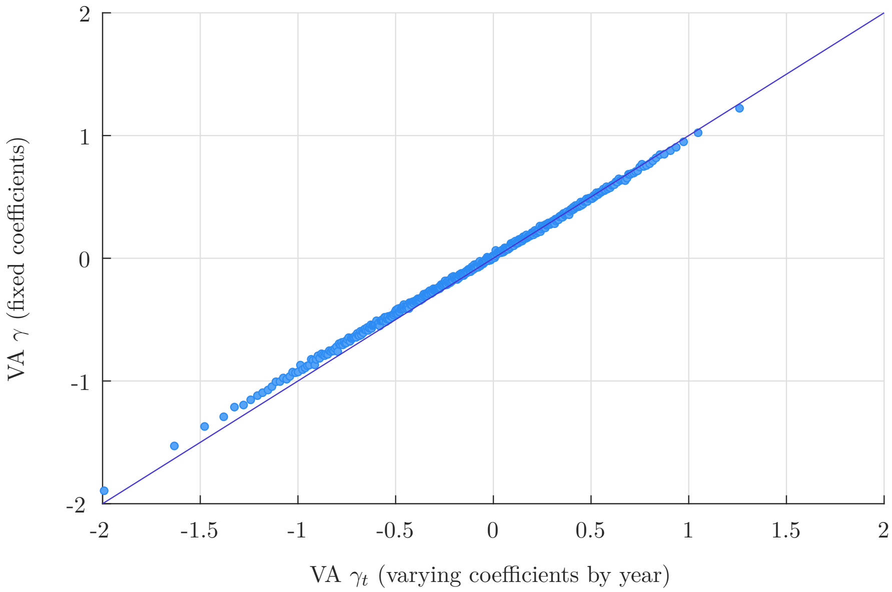
Second, I estimate the production function letting γ vary by year. The correlation between the estimated value added with γ_t and γ fixed is 0.97. A binscatter plot is presented in Figure 1, and it suggests that the assumption regarding γ is not essential to the estimation of the value added as estimates in both cases are very similar.

Table 3: Stability of the production function coefficients

	(1) AVE 2005-2016	(2) AVE 2005-2007 & 2010-2012	(3) AVE 2005-2007	(4) AVE 2005-2011
Mother High School	0.202***	0.208***	0.226***	0.211***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother More than High School	0.276***	0.271***	0.304***	0.279***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother Took PAA	-0.102***	-0.086***	-0.070***	-0.087***
	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.064***	-0.052***	-0.036***	-0.054***
	(0.000)	(0.000)	(0.000)	(0.000)
Parents Married	0.056***	0.072***	0.065***	0.055***
	(0.000)	(0.000)	(0.000)	(0.000)
Single Birth	0.054***	0.058***	0.060***	0.063***
	(0.000)	(0.000)	(0.000)	(0.000)
First Born Child	0.056***	0.074***	0.090***	0.069***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.092***	-0.151***	-0.319***	-0.180***
	(0.000)	(0.000)	(0.000)	(0.000)
Birth Length Group FE	✓	✓	✓	✓
Birth Weight Group FE	✓	✓	✓	✓
Birth Gestation Group FE	✓	✓	✓	✓
Birth Location Group FE	✓	✓	✓	✓
Mother College Exam Math Score Group FE	✓	✓	✓	✓
Mother College Exam Lang Score Group FE	✓	✓	✓	✓
Region Birth FE	✓	✓	✓	✓
Location of Birth FE	✓	✓	✓	✓
School by Year FE	✓		✓	✓
School by Group Year FE		✓		
R^2	0.31	0.28	0.31	0.31
N Obs	2,164,812	1,108,152	563,073	1,282,807

Note: This table shows the regression coefficients of the estimated production function with different subsamples of data. The two first columns show the estimation of value added by school by year fixed effect considering all years, and by school by group of years fixed effects, considering only two periods 2005-2007 and 2010-2012. Columns (3) and (4) repeat these estimations for different subsamples of years. Several large sets of fixed effect groups were ommitted to make the table manageable.

Figure 1: Value Added Estimates with γ_t vs γ



Note: The figure shows a binscatter plot with 100 bins. The X-axis shows school-year value added allowing for γ to vary each year. The Y-axis shows school-year value added fixing γ to not vary each year. Both cases produce estimates of value added that overall are quite similar.

3 Robustness

3.1 Lagged Test Scores

To check for robustness in the estimation of value added, I first present a series of Figures that show the high correspondence between value added estimated with 4th graders' test scores and the same estimation using lagged test scores (test scores in 2nd grade). Figure 2 shows the binscatter plot between the estimated value added (second specification) and the value added estimated with lagged test scores. Figure 3 shows the binscatter plot by type of school (public, private voucher and private non-voucher school), and Figure 4 shows it for profit and non-profit schools.

Figure 2: Value Added and Lagged Values of Test Scores

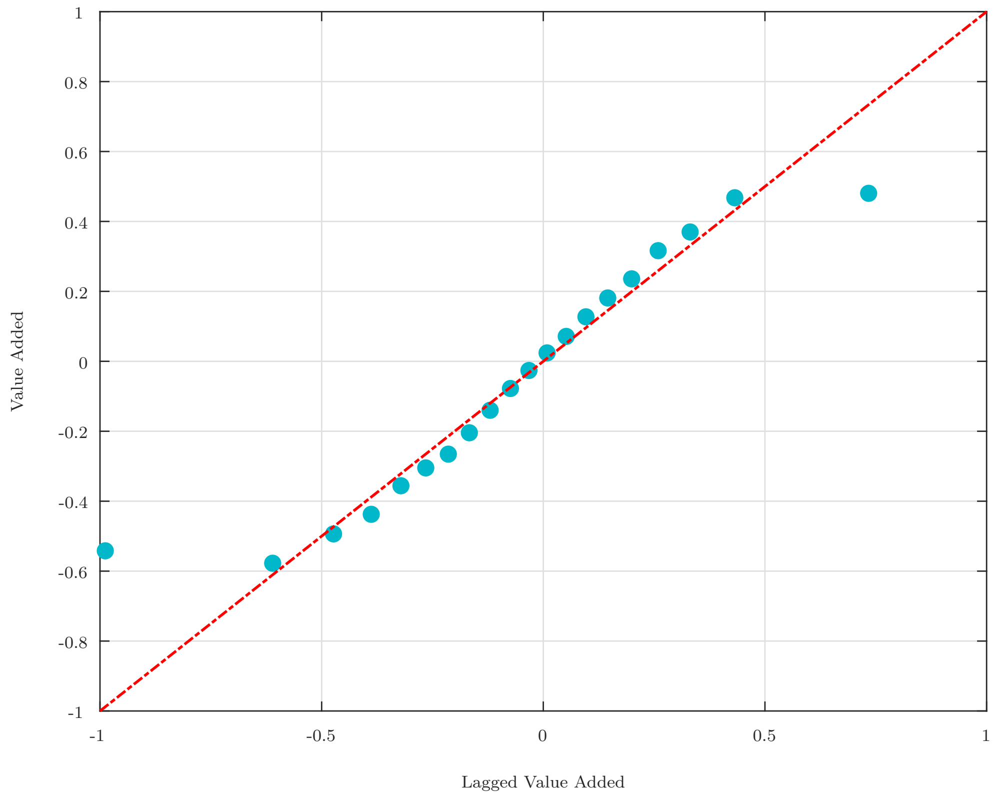
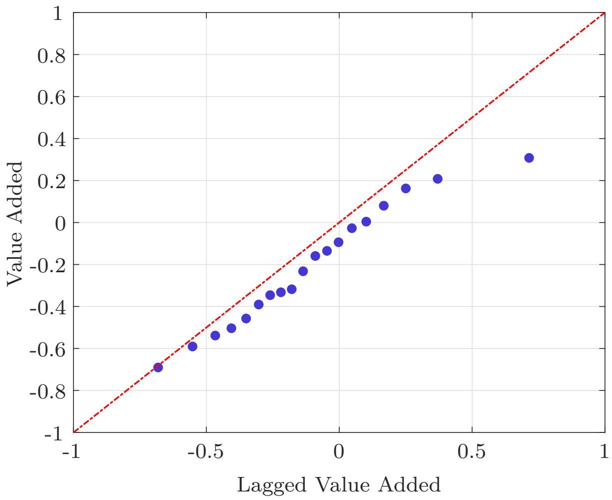
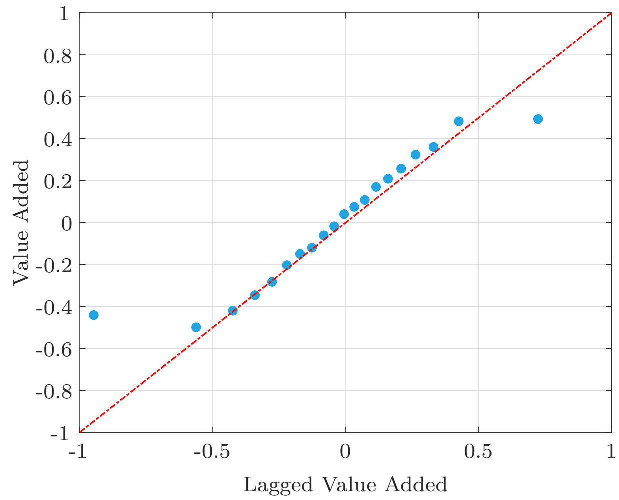


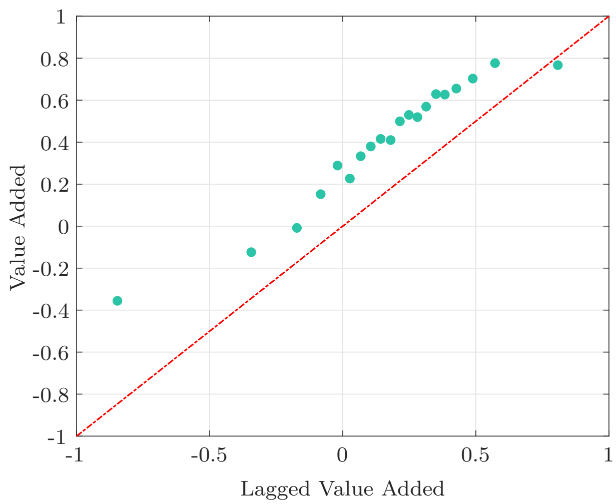
Figure 3: Value Added with and without Pre-Score Information by Administration



(a) Public Schools



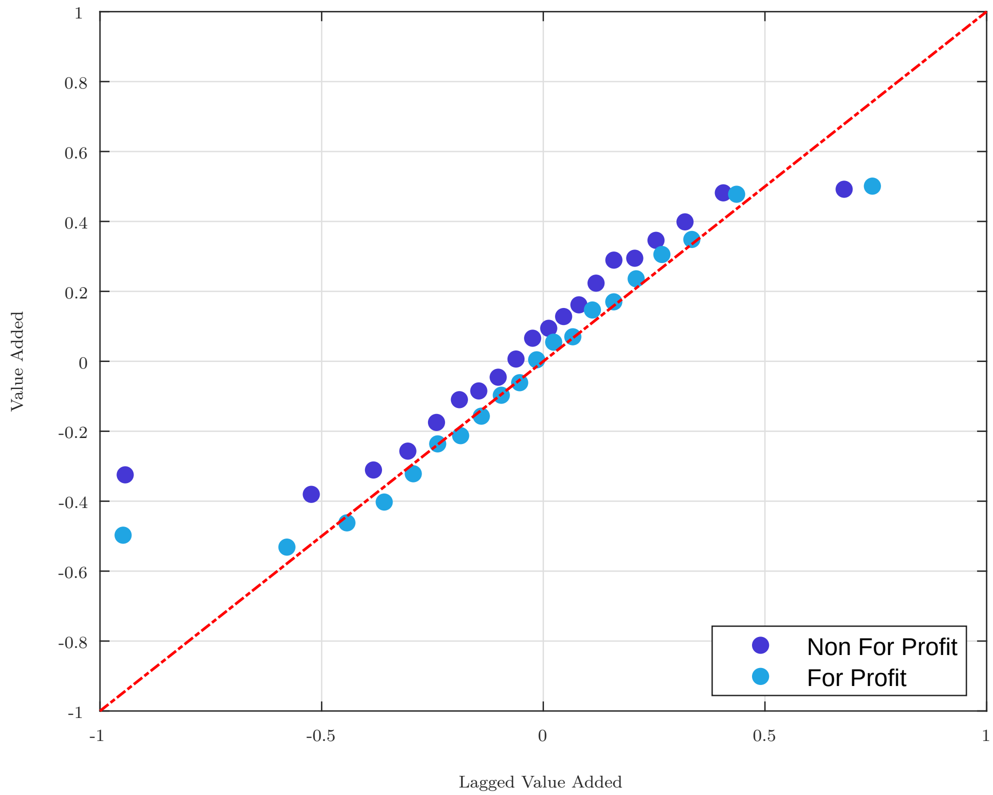
(b) Private Voucher Schools



(c) Private Non-Voucher Schools

Note: This figure replicates Figure 2 by each type of school administration. In this case, we see that the small number of private non-voucher schools seem to have higher value added when pre-scores are not considered. Voucher schools seem to have value added that is similar in both cases.

Figure 4: Value Added and Lagged Values of Test Scores by Profit/Non-Profit Status



I inquire into correlations between current and lagged values for the variables of interest in these estimations. Table 4 shows the correlation by school between the average test score, and the three estimated value added with their respective lagged values. For the estimated value added with the full set of covariates that has school-by-year fixed effects, I estimate the regression coefficient on its lagged value. Results are shown in Table 5.

Table 4: Correlation coefficients with Lagged Values: Average Test Score, VA1, VA2 and VA3

	Avg. Test Score	VA1	VA2	VA3
2005 - 2017	0.353	0.314	0.289	0.366
2005 - 2007	0.361	0.325	0.299	0.355

Note: This table shows the correlation coefficients between the average test score and the value added estimated by three different specifications, with their respective lagged values. Only schools with more than ten students in first grade are considered.

Table 5: Regressions with Lagged Values

	(1) VA2	(2) VA2
Lagged Value Added	0.295***	0.266***
	(0.000)	(0.000)
Constant	-0.122***	-0.087***
	(0.000)	(0.000)
Year FE	✓	✓
Period	2005 - 2007	2005 - 2017
R^2	0.092	0.098
N Obs	5,296	22,410

Note: This table shows the estimated coefficient of the regression of value added over its lagged value, considering the pre-policy period and the complete period. These results are for the two extended estimations of value added, VA2 and VA3 previously defined. Only schools with more than ten students in first grade are considered.

3.2 Estimated q and other school inputs

Another interesting item to see is the relation that estimated quality has with school inputs. Quality of teachers and principals is, in theory, a variable that affects school quality, but this relationship might not be direct or homogeneous between types of schools and school administration.

Figures 5 and 6 show binscatter plots between value added and principals' test scores by type of school (public, private voucher, and private non-voucher) and by profit or non-profit schools. The first figure does not show a marked positive relationship between VA and principals' scores, while the second figure that only considers private schools (voucher or non-voucher) shows a relatively more positive trend.

Figure 7 shows the binscatter plot between value added and teachers' test scores, with a slightly positive trend. Once it is differentiated by type of school, in Figure 8, we can see that both private voucher and private non-voucher schools show a positive relationship with estimated value added, while public schools do not follow this trend. In Figure 9, where only private schools are considered (voucher and non-voucher), the relationship is clearly positive and very close to a 1:1 correspondence.

Figure 5: Value Added and Principal Test Score by Type of School

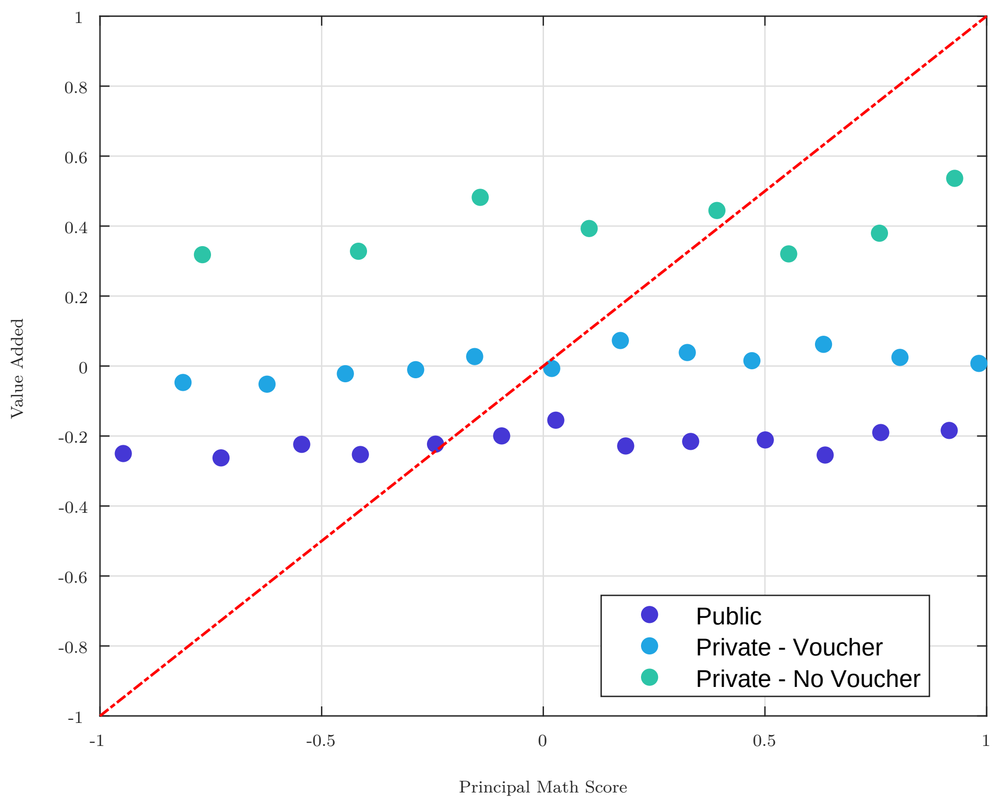


Figure 6: Value Added and Principal Test Score by Profit/Non-Profit Status

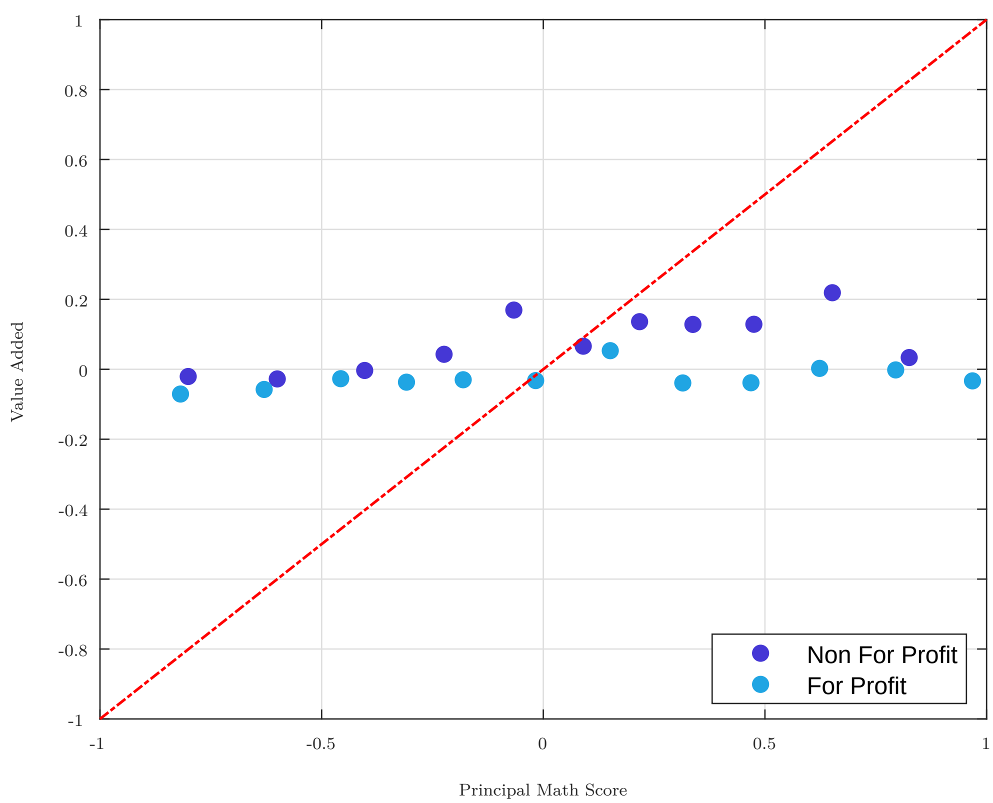


Figure 7: Value Added and Teachers Test Scores

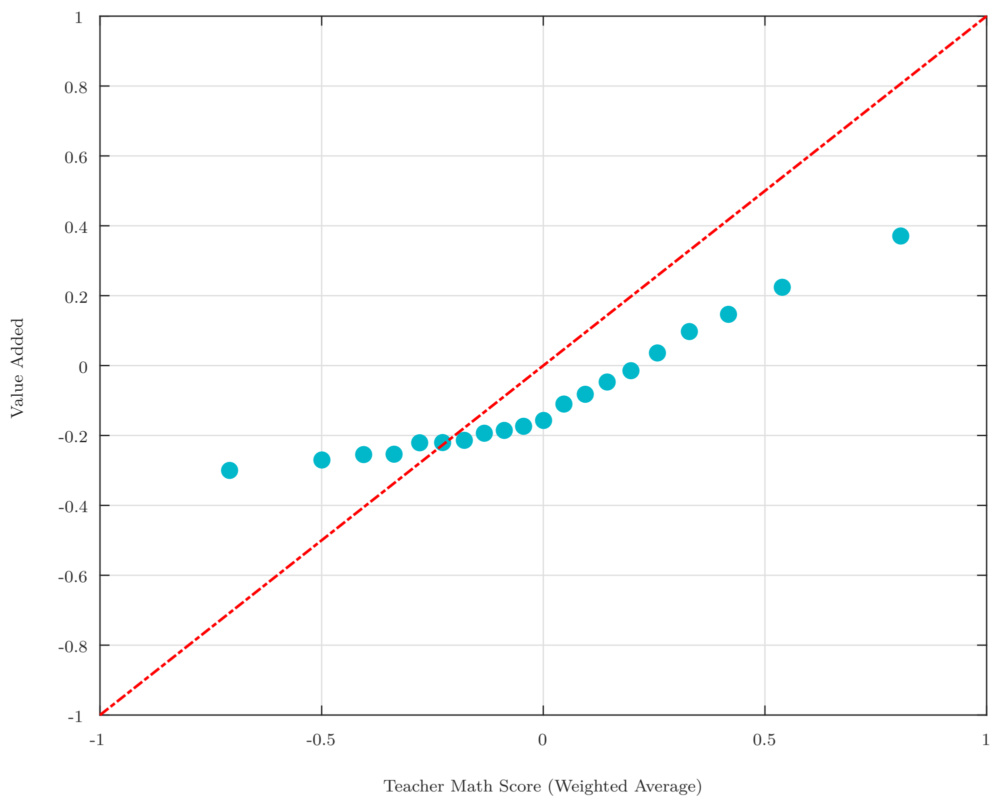


Figure 8: Value Added and Teachers Test Scores by Type of School

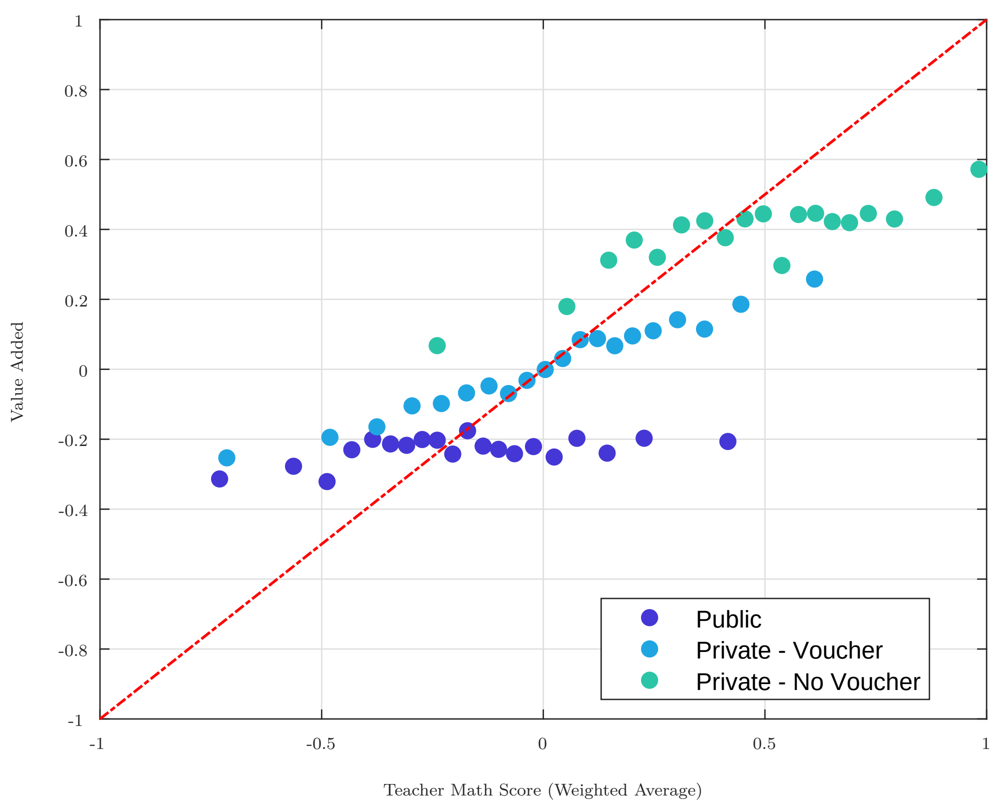
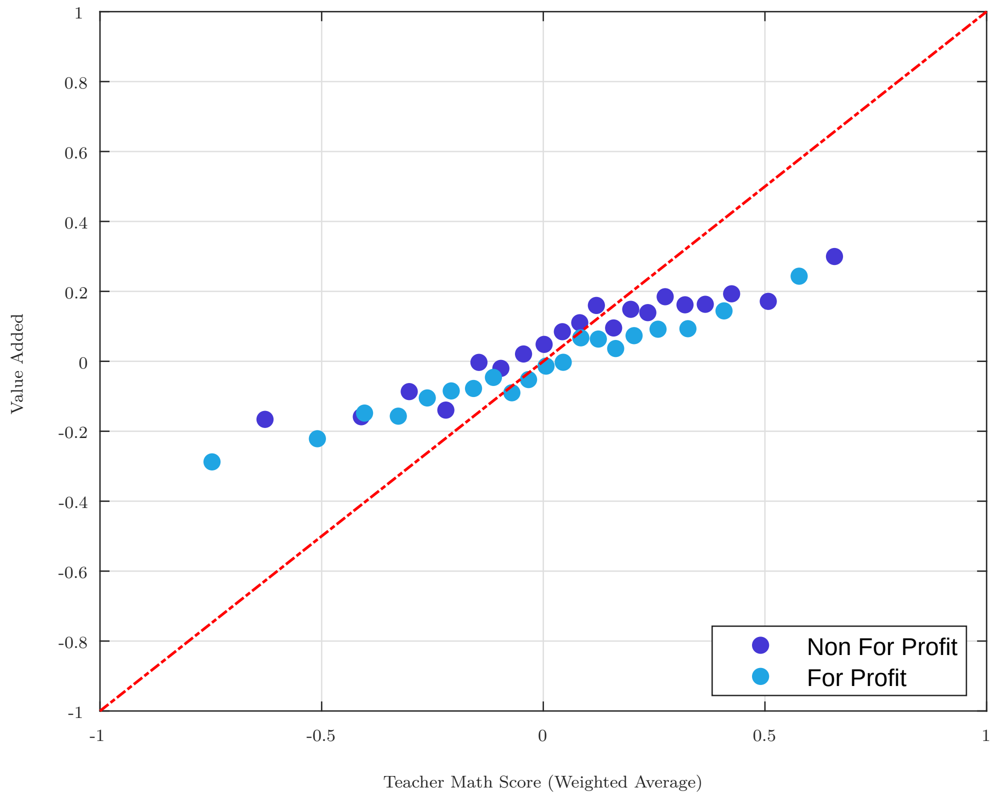


Figure 9: Value Added and Teachers Test Scores by Profit/Non-Profit Status



In the same line, I estimate how these different schools' inputs are related to value added among other measures of quality, adding spending on teachers and revenue per student as well. Table 6 show the results these estimations. The three outputs considered as quality measures are estimated value added, if the school has fines, and if the school has won an academic excellence prize (SNED). The first three columns consider only the years between 2014 and 2017 because of the availability of spending data, while the last three consider all years from 2005 to 2017.

Principals' scores in math have a positive correlation both with estimated value added and with academic excellence prices, while there is no significant correlation with having fines. Principals' scores in language do not present any significant effect. For teachers' scores, the correlation is very similar in math test scores, while language test scores show a positive relationship only with estimated value added. Income per student is positively correlated with value added and with winning and academic prize for years 2005 - 2017, but in recent years, the relationship with the academic prize turns to negative. For the estimation that considers all years, being a traditional school and being religious is positively correlated with quality measures. However, in recent years the relationship loses significance.

The two only variables that are not only positively correlated with quality measures, but negatively related to the bad quality measure (fines) are spending per teacher (not particularly in recent years) and being a for-profit school.

Table 6: Value Added and Spending on Teachers - Market FE

	(1) Quality	(2) Has Fine	(3) Has SNED	(1) Quality	(2) Has Fine	(3) Has SNED
AdminHC Math	0.026***	0.001	0.018***	0.028***	0.002	0.017***
	(0.000)	(0.908)	(0.007)	(0.000)	(0.559)	(0.000)
AdminHC Lang	-0.007	0.006	-0.007	-0.005	0.002	0.003
	(0.311)	(0.367)	(0.359)	(0.251)	(0.520)	(0.518)
TeacherMath WeightedAve	0.282***	0.026	0.102***	0.299***	-0.014	0.154***
	(0.000)	(0.339)	(0.001)	(0.000)	(0.264)	(0.000)
TeacherLang WeightedAve	0.089***	-0.024	0.003	0.101***	-0.013	0.029
	(0.002)	(0.420)	(0.931)	(0.000)	(0.349)	(0.102)
SpendingPerTeacher	0.013***	-0.002***	0.019***			
	(0.000)	(0.001)	(0.000)			
Income per Student	0.041***	-0.001	-0.028**	0.164***	-0.006	0.032***
	(0.000)	(0.957)	(0.028)	(0.000)	(0.348)	(0.000)
Income per Student^2	-0.001***	-0.000	0.001*	-0.004***	-0.000	-0.001***
	(0.001)	(0.731)	(0.092)	(0.000)	(0.971)	(0.001)
Traditional	0.030***	0.011	0.012	0.056***	-0.003	0.089***
	(0.001)	(0.247)	(0.280)	(0.000)	(0.578)	(0.000)
For Profit	0.010	-0.001	0.016	-0.046***	0.022***	-0.096***
	(0.341)	(0.934)	(0.189)	(0.000)	(0.000)	(0.000)
Religious	-0.000	-0.007	0.026**	0.023***	-0.005	0.017***
	(0.970)	(0.465)	(0.023)	(0.000)	(0.358)	(0.009)
Constant	-0.497***	0.217***	-0.208***	-0.351***	0.160***	0.269***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Only 2014 - 2017	x	x	x			
Year FE	x	x	x	x	x	x
Market FE	x	x	x	x	x	x
R^2	0.243	0.039	0.156	0.204	0.133	0.081
N Obs	7,729	7,729	7,729	24,627	24,627	24,627

Note: This table shows the relationship between school characteristics, inputs and the estimated value added. The first three columns include data from 2014 to 2017 which is when school spending data is available. The dependent variables considered are: (i) estimated value added, (ii) if the school has been fined, and (iii) if the school has won a prize for academic excellence (SNED). The independent variables include the average college entrance exam scores for teachers and the school principal. Teacher score averages are weighted by hours employed at the school. Also included is the average spending per classroom teacher and the average school revenue per student. School characteristics include an indicator variable if the school has been open since 1995 and whether the school is for-profit and if it has a religious affiliation. The estimations show that all inputs are positively related to value added and to winning academic excellence prices (except for being a for-profit school). In general, the coefficients are not statistically different from zero in regressions that have "having fines" as the outcome variable.

Finally, it is interesting to see not only if schools' inputs related to quality are correlated with estimated value added, but if families' preferences by schools are correlated too. To analyze this issue, I used data on school applications of applicants in levels preK and 1st grade in 2018. For each applicant, I identify all schools that are within a radius of 1.5km. Then I assign a dummy variable to each school if it was the first preference in the application of the related student. Results are shown in Table 7. The dependent variable illustrates the probability of a neighbor school to be in the first preference of the applicants. As we can see, both the distance between the school and the applicant and the prize that the school charges are negatively correlated with the probability of being her first preference.

On the other hand, value added is positively correlated with being the first school listed. Columns (3) and (4) disaggregate this relationship between quintiles of value added among schools in the neighborhood and add the possibility of a different relation considering the SEP status of the applicant. As we can see, being in the highest quintiles of value added schools in the neighborhood of an applicant improves the probability of being assigned as a first preference, but not in SEP applicants cases.

Table 7: School in First Preference SAE 2018

	(1) School 1st Pref	(2) School 1st Pref	(3) School 1st Pref	(4) School 1st Pref
Distance to School (km)	-0.050*** (0.000)	-0.050*** (0.000)	-0.050*** (0.000)	-0.050*** (0.000)
Ln Avg. Price	-0.005*** (0.000)	-0.005*** (0.000)	-0.002** (0.012)	-0.002** (0.023)
Value Added	0.099*** (0.000)	0.101*** (0.000)		
Value Added * SEP Student		-0.008* (0.070)		
Q2 VA			0.010*** (0.000)	0.011*** (0.001)
Q3 VA			0.039*** (0.000)	0.037*** (0.000)
Q4 VA			0.087*** (0.000)	0.087*** (0.000)
Q5 VA			0.125*** (0.000)	0.129*** (0.000)
Q1 VA * SEP student				0.004 (0.313)
Q2 VA * SEP student				0.001 (0.789)
Q3 VA * SEP student				0.008* (0.067)
Q4 VA * SEP student				0.002 (0.552)
Q5 VA * SEP student				-0.008* (0.054)
Constant	0.124*** (0.000)	0.124*** (0.000)	0.060*** (0.000)	0.057*** (0.000)
R^2	0.033	0.034	0.042	0.042
N Obs	76,216	76,216	76,216	76,216

Note: The dependent variable is constructed as follows: for each applicant in levels preK or 1st grade, all schools that are within a radius of 1.5 km around him are assigned. Then, I built a dummy variable if the school was the first preference in the application of the student. This subsample considers applicants with at least ten schools in the radius predefined. Q1 to Q5 variables correspond to quintiles of Value Added in each student set of neighbors schools.

3.3 Value Added and Exposure to SEP

One of the results shown in *Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students* is that exposure to the policy implies significant positive effects on schools' quality, measured as the estimated value added. Here we show these results in detail.

Schools are categorized in a measure of exposure to the policy based on the concentration of eligible students in the neighborhood. Precisely, it is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school. According to this, I run a difference-in-differences regression, exploiting time and cross-sectional variation, considering schools in the highest and the lowest quintiles of the measure of exposure.

The differences-in-differences model was the following:

$$\hat{q}_{j,t} = \sum_t \text{High Exposure}_j \beta_t + \gamma_t + \varepsilon_{j,t} \tag{2}$$

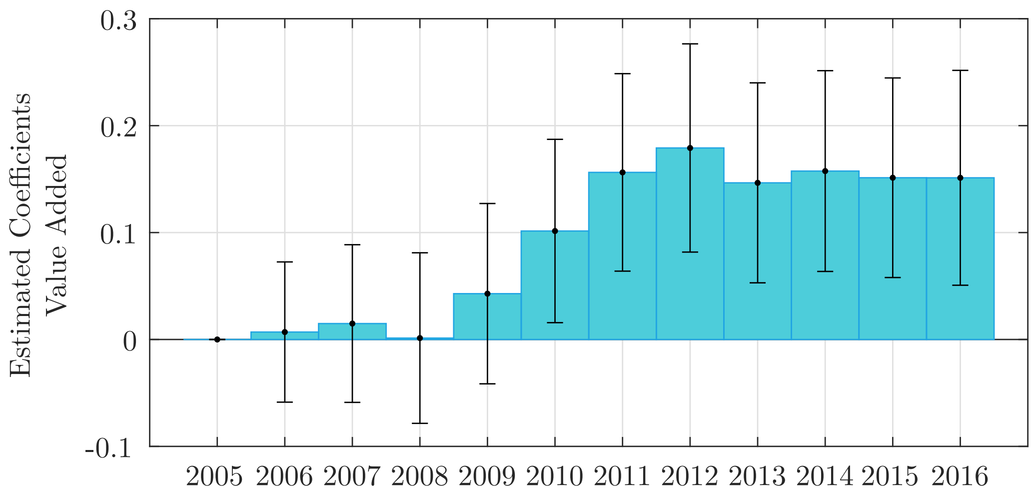
where 2005 serves as the baseline year, and the coefficients γ denotes year fixed effects. The dummy variable Exposure_j takes the value 1 if school j is in the top quintile and 0 if school j is in the bottom quintile of exposure to SEP.

This model is also used to analyze students sorting because of the policy. I perform the same model using fitted test-scores based on students' observables estimated on the pre-policy period ($X_i\gamma$).

Results of the diff-diff model for value added are shown in Figure 10 and in the first column of Table 8. I find that there are no observable pre-trends before SEP is in place, and there are significant effects on school quality in the poorest neighborhoods relative to the richest ones.

Results of the diff-diff model for fitted test-scores are shown in the second column of Table 8. While school value added estimates are large and significant after the policy, estimates for predicted test score index are minimal. This leads us to the conclusion that student characteristics are not changing across schools in different neighborhoods.

Figure 10: Differences in Differences Estimates by Policy Exposure



Note: This figure shows the estimated coefficients from a difference-in-differences estimation on school quality $\widehat{q_{jt}}$ (Value Added). The treatment group correspond to the highest quintile of school level exposure to eligible students, and the control group corresponds to the lowest quintile. The measure of exposure to the policy is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school. The Table 8 in the Appendix show the details of this regression compared to the same regression over predicted test scores $X_i\gamma$ as an index of student characteristics.

Table 8: Differences in Differences Estimates by Policy Exposure

	$\widehat{q_{jt}}$		$X_i\gamma$	
	Coef.	Std.Err	Coef.	Std.Err
Q5 % Poor within 1km (T)	-0.423***	(0.000)	-0.196***	(0.000)
Q5 % Poor within 1km (T) ×2006	0.006	(0.709)	0.001	(0.679)
Q5 % Poor within 1km (T) ×2007	0.015	(0.428)	-0.005**	(0.026)
Q5 % Poor within 1km (T) ×2008	0.001	(0.979)	-0.006**	(0.022)
Q5 % Poor within 1km (T) ×2009	0.040*	(0.065)	0.002	(0.560)
Q5 % Poor within 1km (T) ×2010	0.101***	(0.000)	-0.009***	(0.009)
Q5 % Poor within 1km (T) ×2011	0.155***	(0.000)	-0.007*	(0.074)
Q5 % Poor within 1km (T) ×2012	0.181***	(0.000)	-0.006	(0.141)
Q5 % Poor within 1km (T) ×2013	0.147***	(0.000)	-0.005	(0.200)
Q5 % Poor within 1km (T) ×2014	0.151***	(0.000)	0.016***	(0.000)
Q5 % Poor within 1km (T) ×2015	0.145***	(0.000)	0.001	(0.829)
Q5 % Poor within 1km (T) ×2016	0.145***	(0.000)	0.019***	(0.000)
2006	-0.059***	(0.000)	0.003	(0.220)
2007	-0.056***	(0.000)	0.009***	(0.000)
2008	0.024*	(0.075)	0.002	(0.352)
2009	0.080***	(0.000)	-0.003	(0.296)
2010	0.095***	(0.000)	0.022***	(0.000)
2011	0.060***	(0.000)	0.025***	(0.000)
2012	0.093***	(0.000)	0.025***	(0.000)
2013	-0.000	(0.978)	0.029***	(0.000)
2014	0.019	(0.263)	0.007**	(0.048)
2015	0.059***	(0.000)	0.030***	(0.000)
2016	0.104***	(0.000)	0.016***	(0.000)
Constant	0.186***	(0.000)	0.339***	(0.000)
R^2	0.175		0.400	
N Obs	778899		778899	

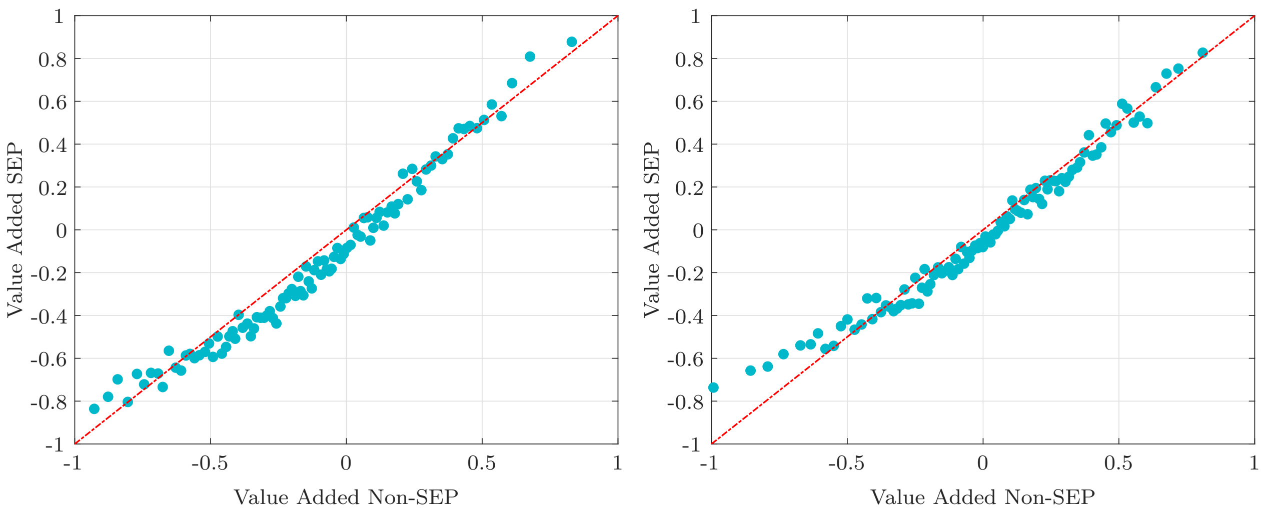
Note: This table shows the estimated coefficients from a difference-in-differences estimation on school quality $\widehat{q_{jt}}$ (Value Added) and the predicted test scores $X_i\gamma$ as an index of student characteristics. The treatment group correspond to the highest quintile of school level exposure to eligible students, and the control group corresponds to the lowest quintile. The measure of exposure to the policy is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school. An extended version of this table is presented in the Online Appendix.

3.3.1 Heterogeneity in Treatment Effects

The evidence presented in earlier sections can raise to the reader some concerns about the possible heterogeneity in the treatment effects. In this line, a relevant robustness check is to show that the estimated school's value added does not differ significantly between priority and non-priority students within a school, and does not differ in growth trends among these groups. Figure 11 shows the relationship between estimated value added within a school for SEP students and Non-SEP students, before and after the SEP policy. The estimation considers fixed effects for combinations of period (pre or post-policy), school and SEP eligibility, and follows the same specification mentioned in previous sections. As we can see, the variables show a direct relationship close to a unitary correlation.

To show that also the growth in value added after the policy remains symmetric between SEP and Non-SEP students, Figure 12 shows the relationship between the increase in estimated school quality for priority and non-priority students within a school. The increase in value added is obtained from the difference between the post-policy and the pre-policy estimated school's value added. to the relationship in absolut value added for pre and post-policy periods, the correlation in growth in value added is close to one.

Figure 11: School Value Added by SEP

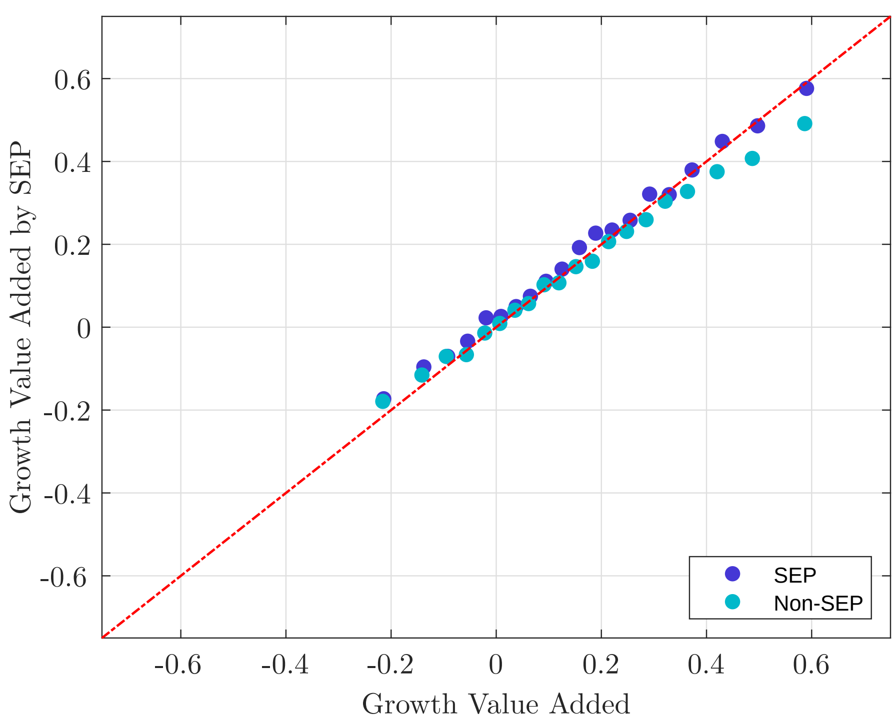


(a) 2005 - 2007

(b) 2010 - 2012

Note: These figures show a bin-scatter where value added considering SEP students are on the X-axis and value added considering Non-SEP students are shown in the Y-axis. The left panel shows the estimation for the pre-policy period and the right panel for the post-policy period.

Figure 12: Value Added growth by Type of Administration



Note: This figure shows the relationship between the growth in estimated school quality from the pre-SEP policy period to the post-SEP policy period, differentiating by type of administration. The pre-policy period accounts for years 2005 to 2007, while the post-policy period for years 2010 to 2012.

3.3.2 Missing data robustness exercise for differences in differences estimates

Cuesta, Gonzalez, and Larroulet (2017) and Feigenberg, Yan, and Rivkin (2019) have pointed out that missing test scores can lead to biased estimates because of selective absences on the day of the test. This issue is relevant for this setting because absenteeism during the test has risen over time, reaching almost 10% of the sample, and the impact of the policy could be confounded with sample selection. However, it could also be less of a concern for this setup since value-added estimates already consider baseline characteristics of students.

For example, a school could manipulate attendance in a way that low-performing students do not go to class on the day of the test, and its results would be inflated. Only if this type of behavior were exacerbated for schools more affected by SEP after the policy is implemented, differences-in-differences estimates would be biased. However, in this setting, value added estimates already consider baseline characteristics, which would prevent an increase in school scores in case advantaged students were kept in the sample.

In any event, this subsection deals with this issue by studying the impact of the policy on missing test scores and re-examining the estimates of the impact of SEP using imputed test-scores. First, I will describe missing data for this setting, and then I will explain the imputation procedure. Finally, I will show how sensible the results are to different imputations.

Missing Data Description

Table 9 and Figure 13 describe the observations used in the differences-in-differences sample. Following Section 6 in the main paper, the universe of observations corresponds to 4th-grade students who are in the selected schooling markets that go to schools either in the two deciles most or least exposed to SEP and that were enrolled in 4th-grade in a school at the end of the year. As described before, the measure of school's exposure to the policy is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school. High exposure is defined as belonging to the top quintile of this measure, and low exposure is defined as belonging to the bottom quintile.

Starting from the raw data set, I drop 7.8% of the sample because of duplicated identifiers or students not enrolled at the end of the year. Students that are sometimes counted twice in the data set are incorrectly registered in their previous school; some other students transfer within the school year. The enrollment registry is linked with the test-score record to keep students in their current schools and drop the observation that corresponds to the previous student enrollment. On the other hand, I drop 2% of observations schools with less than ten scores in any given year, which may lead to scores that are too unreliable. The Quality of Education Agency in Chile also avoids making public results with less than scores for the same reason.

In sum, 9.5% of the raw data set is dropped either because of double-counted students who transferred to other schools, students not enrolled at the end of the year, or students that were in small schools. This number decreases to nearly 8% after 2012 as SIMCE identifiers data quality increases.

I label the rest of the observations are labeled as "usable observations". Nearly 4% of usable observations have missing values in only on the variables used to estimate value added, 7.8% only on test scores, and 1% on both. The next subsection re-estimates the differences-in-differences estimates by imputing the 7% of missing test scores as a robustness check. Value added variables are usually related to administrative records linked with student birth or other information that might not be suspect to manipulation on the day of the test.

Table 9: Sample Decomposition

	Number Obs.	Percentage	Percentage Group
Dropped observations			
Less than 10 observations	17,645	1.8	18.6
Duplicated Identifier	77,125	7.8	81.4
Total dropped	94,770	9.5	100.0
Usable Observations			
Complete observations	778,899	78.4	86.6
Only Missing variables VA Model	34,695	3.5	3.9
Missing variables VA Model & Test scores	7,835	0.8	0.9
Only Missing Test-scores	77,587	7.8	8.6
Total usable	899,016	90.5	100.0
Total	993,786	100	

Note: This table describe the sample decomposition of the 4th grade students sample considered in markets.

Figure 13 shows the evolution of usable observations over-time. The number of missing test scores increases over-time. There seems to be a faster increase of missing test scores on schools that are more exposed to SEP, which is the main concern of this section.

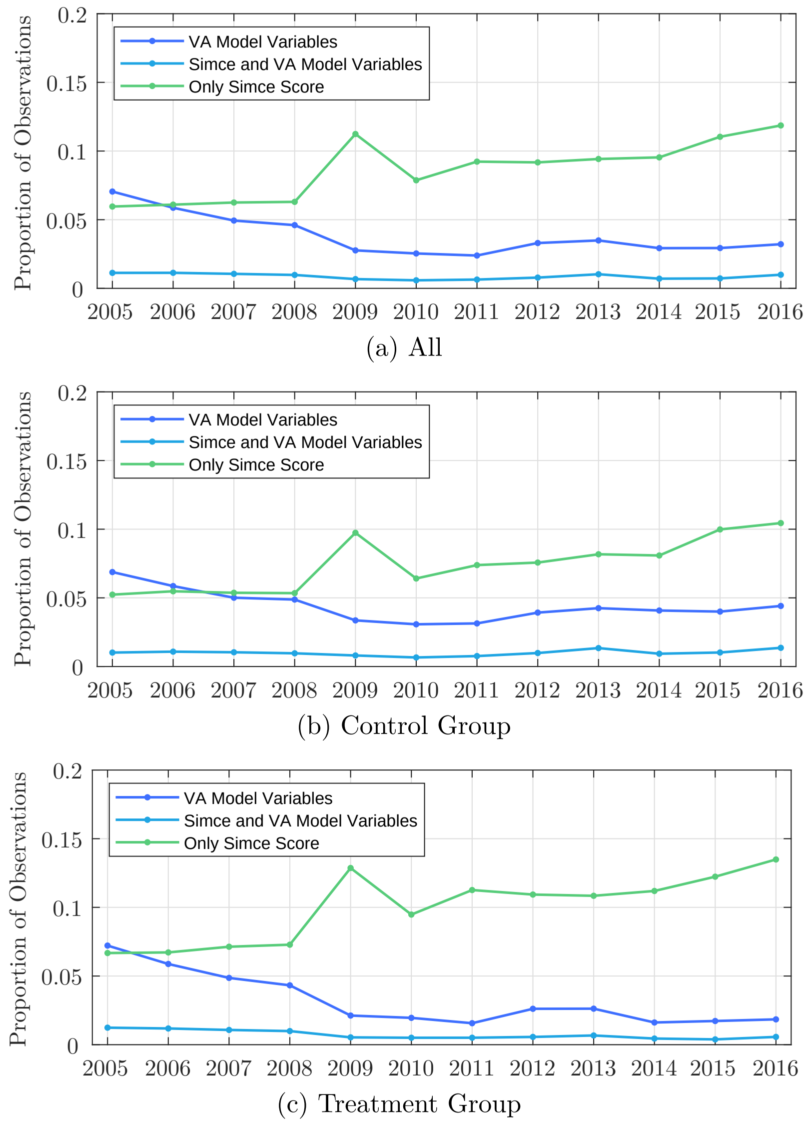
Imputation Procedure

The missing test scores imputation procedure closely follows Cuesta, Gonzalez, and Larroulet (2017), and it includes both excused and non-excused missing records. For each school separately, I regress the test score on a set of yearly dummies and GPA, GPA squared, an indicator of whether students were in fourth grade last year, and an indicator of whether students were in the same school last year. Then, I use that regression to predict test scores for absent students and then estimate the value added model using observed and imputed scores. To account for the uncertainty of the estimates, I draw 100 parameters from the asymptotic distribution from each school. This procedure allows for estimating 100 imputations for each missing score in each school. I pool these estimates into three different imputation measures. The first one averages all the imputations, the second one averages the lowest 25 imputations, and the last one averages the highest 25 imputations.

To check the imputation model, I use the same cross-validation procedure from Cuesta, Gonzalez, and Larroulet (2017). First, I delete ten percent of the lowest GPA scores within each school year. Second, I run each school regression without those observations. Third, I draw 100 imputations for all missing data, including these new missing data. Last, I compare the imputed data against the real data. Figure 14 shows binscatter plots of true test scores against imputed scores. On average, we can see that the imputations match the true scores, which validates the use of the imputation model for this setting[1].

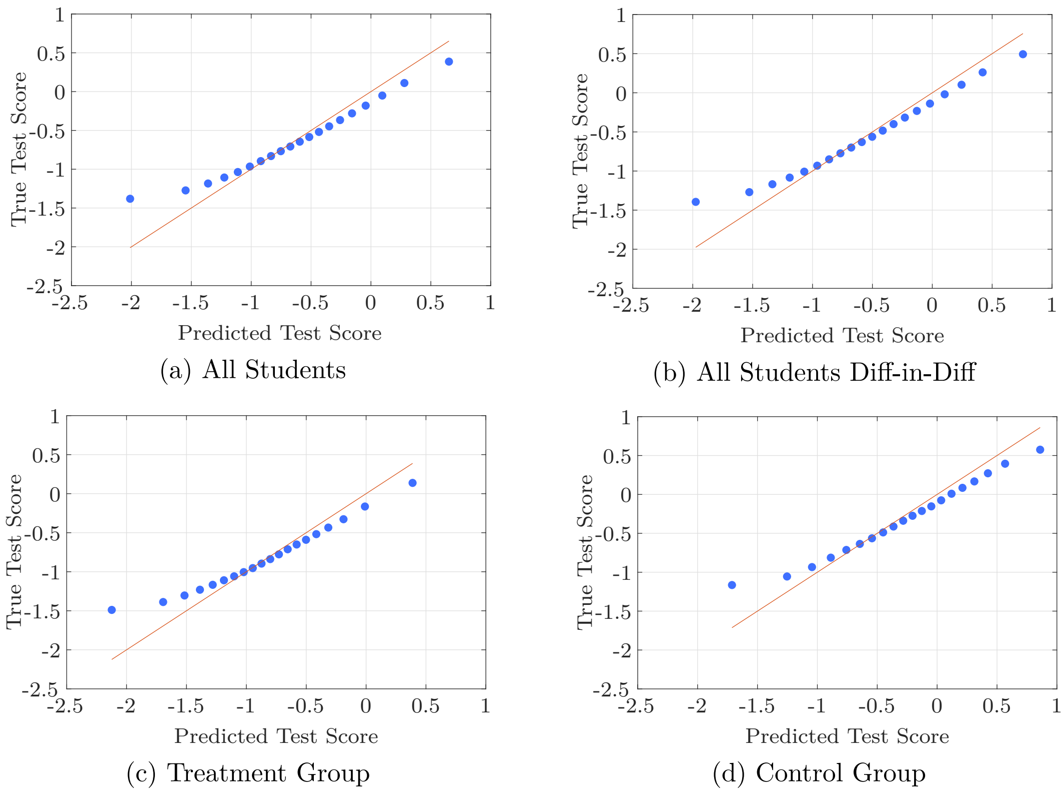
[1] I observe some discrepancies for the lowest values. Imputations turn out to be smaller than the actual scores at the very bottom of the distribution. However, if anything, selective attendance would be more visible because a bad GPA is assigned a worse imputation than its real score. Also, there does not seem to be much difference between the Treatment and Control group, which makes this issue less relevant for causal inference.

Figure 13: Evolution of missing observations across time



Note: This figure shows the evolution of usable observations over-time. Panel (a) shows the evolution considering all the sample, while panels (b) and (c) consider only control and treatment groups, respectively. The treatment group is defined by belonging to the top quintile of the measure of school's exposure to the policy, while control group is defined by belonging to the bottom quintile. The measure of school's exposure to the policy is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school.

Figure 14: Imputation Model Check



(a) All Students

(b) All Students Diff-in-Diff

(c) Treatment Group

(d) Control Group

Note: These figures show binscatter plots of true test scores (y-axis) and predicted test scores (x-axis). Predicted test score are observations that were dropped randomly following the Cross-Validatrion procedure from Cuesta, Gonzalez, and Larroulet (2017). The red line is the Y = X line. Panels (b), (c) and (d) restrict the model to the universe of students considered in the Differences-in-Differences model from the main paper. Panels (c) and (d) consider only the treatment and control group, respectively. The treatment group is defined by belonging to the top quintile of the measure of school's exposure to the policy, while control group is defined by belonging to the bottom quintile. The measure of school's exposure to the policy is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school.

Differences-in-differences results using imputed test scores

Once the missing scores are imputed, and the value added model is run using imputed scores, I incorporate these new value added estimates to the analysis and study how robust is the impact of SEP to imputations. The results of this exercise are displayed in Table 10. The first column shows the impact of the policy on missing scores. Columns (2) to (5) show the effect of policy exposure on value added, similar to the first column in Table 8. The second column shows the results without considering imputations, which are the same estimates presented before and in the main paper. The third, fourth, and fifth columns consider imputations in three different ways: using value added models where missing scores are imputed with the average of the lowest 25 imputations, using all 100 different imputations, and using the highest 25 imputations.

The first column of Table 10 shows some positive and significant coefficient for the early years of SEP from 2010 to 2012 but doesn't seem to affect the results significantly. As expected, poorer schools have a higher rate of missing data. After including imputations, columns three to five show similar results to column two, reassuring that differences-in-differences are robust to considering imputations. Missing data play a role, although it is modest: differences are at the second decimal. In general, using the highest imputation values result in a higher SEP impact and using the lowest imputation in a lower SEP impact. The most significant differences between no imputation and imputation are observed between 2010 and 2012, which is the same years in which we see positive coefficients on the first column. Once imputations are incorporated, the estimates of the effect of SEP on test scores get flatter and closer to 0.14.

Table 10: Differences-in-differences estimates on Missings Data and Value Added with and without imputations

	(1)	(2)	(3)	(4)	(5)
	Missing	No Imputations		Imputations	
			Lowest 25	All	Highest 25
Q5 % Poor within 1km (T)	0.017***	-0.423***	-0.443***	-0.442***	-0.441***
	(0.003)	(0.026)	(0.026)	(0.026)	(0.026)
Q5 % Poor within 1km (T) ×2006	-0.003	0.006	0.009	0.009	0.009
	(0.004)	(0.017)	(0.017)	(0.017)	(0.017)
Q5 % Poor within 1km (T) ×2007	0.001	0.015	0.020	0.021	0.021
	(0.005)	(0.019)	(0.019)	(0.019)	(0.019)
Q5 % Poor within 1km (T) ×2008	0.003	0.001	0.003	0.003	0.004
	(0.004)	(0.020)	(0.020)	(0.020)	(0.020)
Q5 % Poor within 1km (T) ×2009	0.012	0.040*	0.033	0.038*	0.042**
	(0.010)	(0.022)	(0.022)	(0.021)	(0.021)
Q5 % Poor within 1km (T) ×2010	0.012***	0.101***	0.091***	0.094***	0.097***
	(0.004)	(0.022)	(0.022)	(0.022)	(0.022)
Q5 % Poor within 1km (T) ×2011	0.020***	0.155***	0.145***	0.149***	0.154***
	(0.005)	(0.024)	(0.023)	(0.023)	(0.023)
Q5 % Poor within 1km (T) ×2012	0.013***	0.181***	0.169***	0.173***	0.177***
	(0.005)	(0.025)	(0.025)	(0.025)	(0.025)
Q5 % Poor within 1km (T) ×2013	0.003	0.147***	0.143***	0.145***	0.147***
	(0.005)	(0.024)	(0.024)	(0.024)	(0.024)
Q5 % Poor within 1km (T) ×2014	0.010*	0.151***	0.146***	0.149***	0.153***
	(0.006)	(0.024)	(0.024)	(0.024)	(0.024)
Q5 % Poor within 1km (T) ×2015	-0.000	0.145***	0.136***	0.137***	0.137***
	(0.005)	(0.024)	(0.024)	(0.024)	(0.024)
Q5 % Poor within 1km (T) ×2016	0.006	0.145***	0.132***	0.135***	0.137***
	(0.006)	(0.025)	(0.025)	(0.025)	(0.025)
Constant	0.062***	0.186***	0.213***	0.205***	0.197***
	(0.002)	(0.019)	(0.019)	(0.019)	(0.019)
Year FE	✓	✓	✓	✓	✓
R^2	0.006	0.175	0.177	0.180	0.183
N Obs	899,016	778,899	856,486	856,486	856,486

Note: This table shows the estimation of a differences-in-differences methodology of missing test scores and value added over the policy exposure, considering imputations excercises. The treatment group corresponds to the highest quintile of school level exposure to eligible students, and the control group corresponds to the lowest quintile. The measure of exposure to the policy is calculated as the share of SEP eligible students that live within a 1.5 km radius from the school. Column (1) shows the impact of the policy on missing scores considering both excused and non-excused. Columns (2) to (5) show the effect of policy exposure on value added. Column (2) shows the results without considering imputations, similar to the first column in Table 8. Columns (3) to (5) consider imputations using value added models where missing scores are imputed with the average of the lowest 25 imputations, using all 100 different imputations, and using the highest 25 imputations.

Feigenberg, Yan, and Rivkin (2019) states that part of the impact of SEP on test scores dissipates when considering missing data. However, there are several reasons why these findings are different. First, as already mentioned, I use value added measures and not test scores. Second, the design considers externalities across schools in the spirit of Correa, Parro, and Reyes (2014).

Control schools might also be affected by the treatment because of competitive pressure, which is why the sample only considers schools more and less exposed to SEP. Third, this missing data robustness exercise considers treatment and control schools are defined by exposure to SEP and not a noisy proxy as being a public school.

References

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(4), 351–398.

Cuesta, J. I., F. Gonzalez, and C. Larroulet (2017): "Distorted quality signals in school markets," Manuscript.

Feigenberg, B., R. Yan, and S. Rivkin (2019): "Illusory Gains from Chile's Targeted School Voucher Experiment*," *The Economic Journal*.