

How can progressive vouchers help the poor benefit from school choice? Evidence from the Chilean voucher system

Josefa Aguirre

Online Appendix 1: School characteristics in 2007

This section presents average school characteristics in 2007 for schools that in 2012 were: public; private voucher and not charging add-ons to parents; private voucher charging add-ons to parents, and had joined the program; private voucher, charging add-ons to parents, and had not joined the program.

Table A.2:
School Characteristics in 2007

	(1) Test Scores	(2) SES	(3) Add-On	(4) Size	(5) Class Size
Public	235.8 (27.5)	9.4 (1.8)	0.0 (0.0)	26.5 (30.0)	17.4 (13.3)
Private Voucher w/ No Add-On that joined the program	231.7 (29.0)	9.4 (2.2)	0.0 (0.0)	25.9 (29.8)	18.3 (14.6)
Private Voucher w/ Add- On that joined the program	256.0 (22.9)	12.3 (1.2)	41.8 (34.2)	54.5 (37.9)	32.7 (9.9)
Private Voucher w/ Add- On that did not join the program	264 (22.7)	13.1 (1.2)	69.4 (45.9)	57.2 (43.0)	31.0 (9.3)

Notes: Table A.2 presents average characteristics in 2007 for different types of subsidized schools in 2012. Test score equals the average result on the fourth grade standardized test in 2007; SES equals the average years of education of mothers of students attending those schools; add-on equals the total amount charged to noneligible parents in those schools in USD; school size equals the cohort size at those schools; class size equals the average class size at those school.

Online Appendix 2: Students with Address Information

This section shows how students for whom I obtain exact addresses (40%) compare to the general sample of students used in the analysis. As can be seen, students with address information have slightly higher SES characteristics, but are not substantially different from the general sample used in the analysis.

Table A.3:
Characteristics of Students with Address Information

	Sample	Students with address information
Mother's Education	10.554 (2.889)	10.906 (2.742)
Father's Education	10.345 (3.001)	10.722 (2.936)
Income USD	600.330 (508.713)	660.985 (512.251)
No. of Books	24.393 (24.201)	27.328 (25.221)
Internet	0.477 (0.499)	0.564 (0.496)
Computer	0.675 (0.469)	0.736 (0.441)
Attended Childcare (0–2)	0.202 (0.402)	0.219 (0.413)
Obs	31,641	11,961

Notes: Table A.3 shows how average characteristics of students with address information compare to average characteristics of students in the sample. Standard deviations in parentheses.

Online Appendix 3: Estimates with Alternative Bandwidths

This section presents the main results from this study using alternative bandwidths. Results from Table 3 use a bandwidth of 1200 points. In this section, I present estimates using a 500, 1000, 1500, and 2000 bandwidths. All estimates are for the effect of being eligible for progressive vouchers, where the discontinuity is used as an instrument for eligibility.

Table A.4:
School Choice and Educational Outcomes with Alternative Bandwidths

	(1) School Private	(2) School Test Scores	(3) School SES	(4) School Add-On	(5) School Class Size	(6) School Distance (Miles)	(7) Student Language Second Grade	(8) Student Language Fourth Grade	(9) Student Math Fourth Grade
Panel A: 500 Bandwidth									
Eligible	0.029 (0.026)	0.047 (0.039)	0.045 (0.034)	4.039** (1.689)	0.142 (0.553)	0.024 (0.151)	-0.037 (0.061)	-0.013 (0.056)	-0.102** (0.052)
Mean Control	0.478	0.444	0.653	12.746	29.559	1.299	-0.188	0.085	0.157
Observations	13,474	13,064	13,042	13,412	13,263	5,149	10,550	9,944	9,998
Panel B: 1000 Bandwidth									
Eligible	0.009 (0.018)	0.014 (0.027)	0.033 (0.024)	3.012** (1.183)	0.094 (0.389)	0.048 (0.101)	-0.040 (0.042)	-0.003 (0.039)	-0.048 (0.036)
Mean Control	0.494	0.456	0.680	14.291	29.602	1.339	-0.166	0.084	0.138
Observations	26,360	25,544	25,503	26,244	25,936	9,979	20,690	19,609	19,698
Panel C: 1500 Bandwidth									
Eligible	0.001 (0.015)	0.001 (0.022)	0.021 (0.019)	2.819*** (0.975)	-0.210 (0.319)	0.028 (0.082)	-0.038 (0.035)	-0.022 (0.033)	-0.038 (0.030)
Mean Control	0.505	0.469	0.706	14.872	30.124	1.382	-0.160	0.093	0.147
Observations	39,788	38,541	38,484	39,607	39,157	14,984	31,223	29,633	29,709
Panel D: 2000 Bandwidth									
Eligible	-0.003 (0.013)	-0.004 (0.019)	0.013 (0.017)	2.648*** (0.851)	-0.317 (0.279)	0.016 (0.070)	-0.024 (0.030)	-0.019 (0.028)	-0.033 (0.026)
Mean Control	0.500	0.474	0.703	14.761	30.042	1.339	-0.168	0.084	0.131
Observations	54,702	52,950	52,867	54,457	53,803	22,174	42,955	40,782	40,838

Notes: Table A.4 shows results using alternative bandwidths. All estimates include controls for mother's education, father's education, and region. All cells contain instrumental variable estimates, where the discontinuity is used as an instrument for being eligible for progressive vouchers. Results from rdrobust (Calonico, Cattaneo, and Titiuk 2014). Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Online Appendix 4: Compliers' Characteristics

This section uses Abadie (2002) to look at the characteristics of compliers in the sample, meaning students who become eligible for the extra voucher amount because their socioeconomic ranking is below the threshold for program eligibility, but who would not be eligible for the program otherwise. The Table compares the characteristics of compliers to those of all students who are close to the threshold.

Table A.5:
Compliers' Characteristics

	Sample	Compliers
Mother's Education	10.553	10.786
Father's Education	10.344	10.594
Income USD	600.365	624.379
No. of Books	24,393	26.037
Internet	0.477	0.514
Computer	0.674	0.700
Attended Childcare (0-2)	0.202	0.230

Notes: Table A.5 shows how average characteristics of compliers compare to average characteristics of students who are close to the threshold (Abadie 2002).

Online Appendix 5: Effect of the Program on Schools Attended by Students in Fourth Grade

This section estimates the effects of the program on the characteristics of the schools students attend in fourth grade.

Table A.6:
Characteristics of the Schools Attended in Fourth Grade

	(1) School Private	(2) School Test Scores	(3) School SES	(4) School Add- On	(5) School Class Size	(6) School Distance (Miles)
Panel A: Impact of Being Below the Cutoff (Reduced Form)						
$R \leq Cutoff$	0.018 (0.012)	-0.002 (0.017)	0.005 (0.015)	2.124*** (0.755)	0.004 (0.250)	0.010 (0.223)
Mean Control	0.488	0.481	0.712	15.912	29.848	2.557
Observations	30,982	30,242	30,191	30,837	30,479	11,952
Panel B: Impact of Being Eligible for Targeted Vouchers (IV)						
$R \leq Cutoff$	0.026 (0.017)	-0.003 (0.025)	0.007 (0.021)	3.055*** (1.088)	0.006 (0.360)	0.013 (0.311)
Mean Control	0.482	0.485	0.710	15.096	29.917	2.555
Observations	30,982	30,242	30,191	30,837	30,479	11,952

Notes: Table A.6 shows results on school choices as of fourth grade. All estimates include controls for mother's education, father's education, and region. Panel A contains reduced-form estimates, and Panel B contains instrumental variable estimates, where the discontinuity is used as an instrument for being eligible for progressive vouchers in first grade. Results from rdrobust (Calonico, Cattaneo, and Titiuk 2014). Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

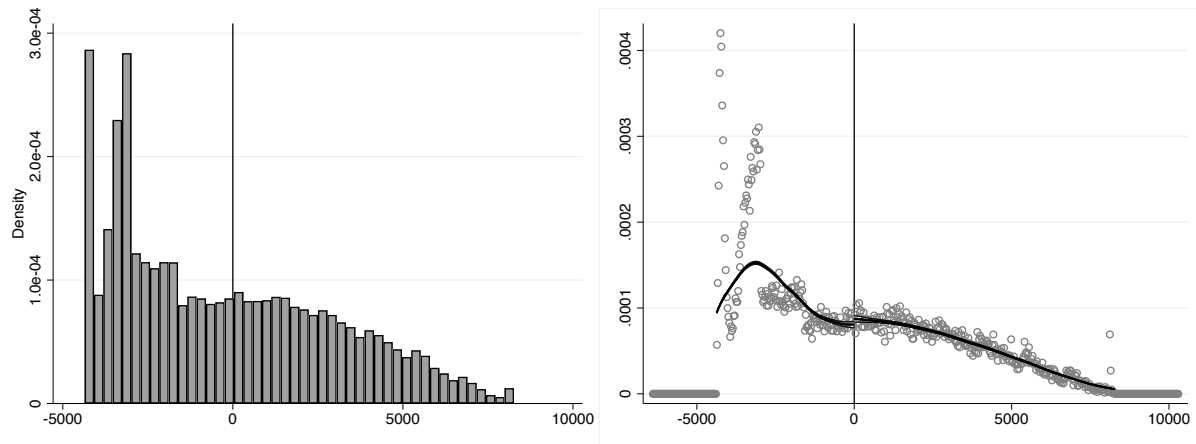
Online Appendix 6: Robustness Checks for Heterogeneous Results

This section performs standard robustness checks for the different subgroups analyzed in the study: (1) students whose mother has less than a high school education, (2) students whose mother has more than a high school education, and (3) students living in neighborhoods where there was at least one private voucher schools that charged add-ons and joined the program that outperformed all the other public and private voucher schools that charge no add-ons within a 0.7 miles radius, and (4) students living in neighborhoods where there was no private voucher schools that charged add-ons and joined the program that outperformed all the other public and private voucher schools that charge no add-ons within a 0.7 miles radius.

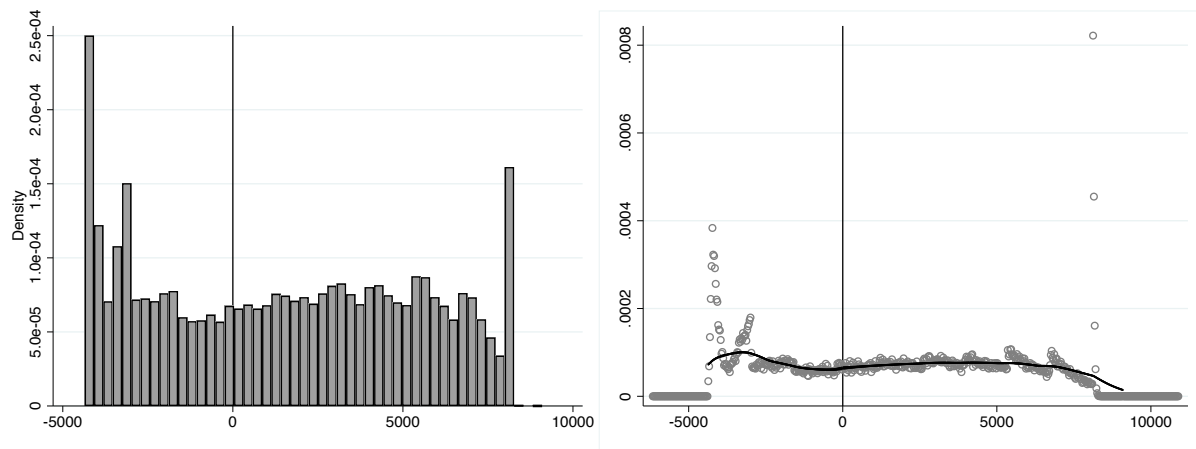
Table A.7:
Robustness Check: Heterogeneous Effects by Mothers' Education

	(1) Mother's Education	(2) Father's Education	(3) Income US	(4) Books	(5) Internet	(6) Computer	(7) Attended Childcare (0-2)
Panel A: Mother Has Less Than High School Education							
Eligible	-0.021 (0.097)	-0.017 (0.127)	-1.144 (12.792)	1.113 (0.833)	0.005 (0.022)	-0.022 (0.022)	0.030* (0.017)
Mean Control	7.873	8.736	419.627	18.006	0.331	0.531	0.164
Observations	9,999	9,463	9,896	9,870	8,326	9,550	9,789
Panel B: Mother Has High School Education or Tertiary Education							
Eligible	-0.046 (0.042)	0.044 (0.099)	-28.660 (22.771)	-0.461 (0.993)	-0.006 (0.020)	0.012 (0.016)	0.009 (0.016)
Mean Control	12.660	11.528	750.040	29.443	0.603	0.789	0.224
Observations	12,942	12,295	12,843	12,794	11,852	12,513	12,814

Notes: Table A.7 shows that in both subsamples, students around the cutoff are similar in terms of their socioeconomic characteristics. Results from rdrobust (Calonico, Cattaneo, and Titiuk 2014). Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.



(A) Mother Has Less Than High School Education



(B) Mother Has High School Education or Tertiary Education

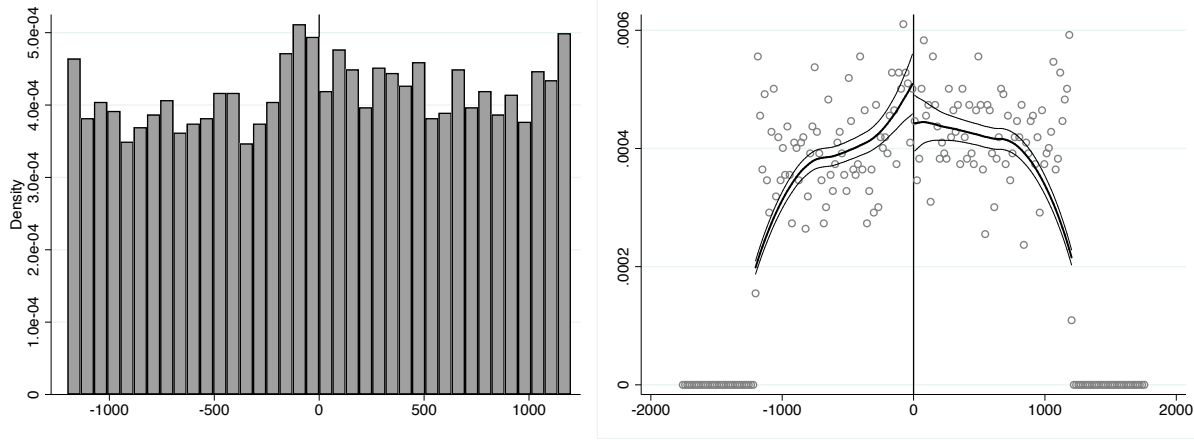
Figure A.2: Visual Evaluation of Robustness Checks: Heterogeneous Effects by Mother's Education

Figure A.2 shows that there is no evidence of socioeconomic manipulation across the threshold in either group. In each case, the histogram shows scores relative to the cutoff value and results of the McCrary (2008) test. Figure (A): $r(\text{bandwidth}) = 2012.909$ $r(\text{binsize}) = 28.250$ $r(\text{se}) = 0.034$ $r(\text{theta}) = 0.077$. Figure (B): $r(\text{bandwidth}) = 1789.870$ $r(\text{binsize}) = 26.85$ $r(\text{se}) = 0.031$ $r(\text{theta}) = 0.007$.

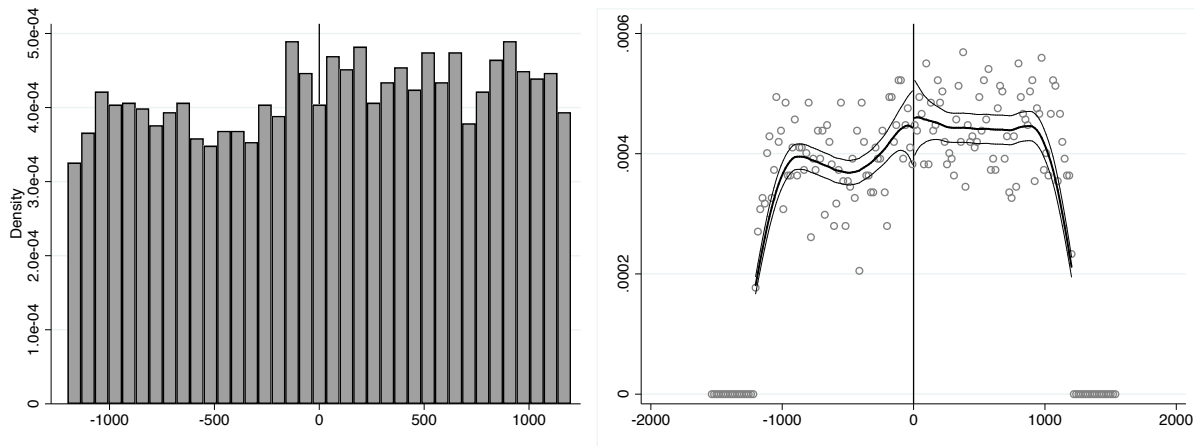
Table A.8:
Robustness Check: Heterogeneous Effects by Neighborhood

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mother's Education	Father's Education	Income US	Books	Internet	Computer	Attended Childcare (0-2)
Panel A: Policy Altered Choice Sets Significantly in a 0.7 Mile Radius							
Eligible	0.124 (0.168)	-0.016 (0.190)	-31.071 (34.422)	-0.014 (1.629)	-0.042 (0.033)	-0.026 (0.028)	0.024 (0.027)
Mean Control	10.975	10.865	696.092	27.898	0.609	0.737	0.209
Observations	4,439	4,249	4,456	4,438	4,045	4,321	4,429
Panel B: Policy Did Not Alter Choice Sets Significantly in a 0.7 mile Radius							
Eligible	-0.018 (0.183)	0.033 (0.201)	-3.646 (35.601)	-0.370 (1.621)	-0.056* (0.034)	0.014 (0.029)	0.013 (0.027)
Mean Control	10.816	10.481	656.391	26.729	0.533	0.725	0.215
Observations	4,339	4,156	4,360	4,347	3,976	4,223	4,326

Notes: Table A.8 shows that in both subsamples, students around the cutoff are similar in terms of their socioeconomic characteristics. Results from rdrobust (Calonico, Cattaneo, and Titiuk 2014). Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.



(A) Policy Altered Choice Sets Significantly in a 0.7 Mile Radius



(B) Policy Did Not Alter Choice Sets Significantly in a 0.7 Mile Radius

Figure A.3: Visual Evaluation of Robustness Checks: Heterogeneous Effects by Neighborhood

Figure A.3 shows that there is no evidence of socioeconomic manipulation across the threshold in either group. In each case, the histogram shows scores relative to the cutoff value and results of the McCrary (2008) test. Figure (A): $r(\text{bandwidth}) = 539.037$ $r(\text{binsize}) = 17.307$ $r(\text{se}) = 0.077$ $r(\text{theta}) = -0.147$. Figure (B): $r(\text{bandwidth}) = 316.02$ $r(\text{binsize}) = 17.56$ $r(\text{se}) = 0.105$ $r(\text{theta}) = 0.035$.

Online Appendix 7: Conditional Independence Tests

This section presents results for the conditional independence test required to implement the Angrist and Rokkanen (2015) approach in section V.E.4. Table A.9, Panels A and C show the relationship between the running variable and the outcomes of interest for students who are below and above the cutoff. The running variable is divided by 1000, meaning that, for example, an increase of 1000 points in the socioeconomic score predicts a 0.014 standard deviation increase in the test scores of chosen schools for students who are below the threshold. Next, Panels B and D show this relationship once I control by the set of socioeconomic indicators. As can be seen in Panels B and D, including the set of controls significantly reduces the relationship between the running variable and the outcomes of interest.

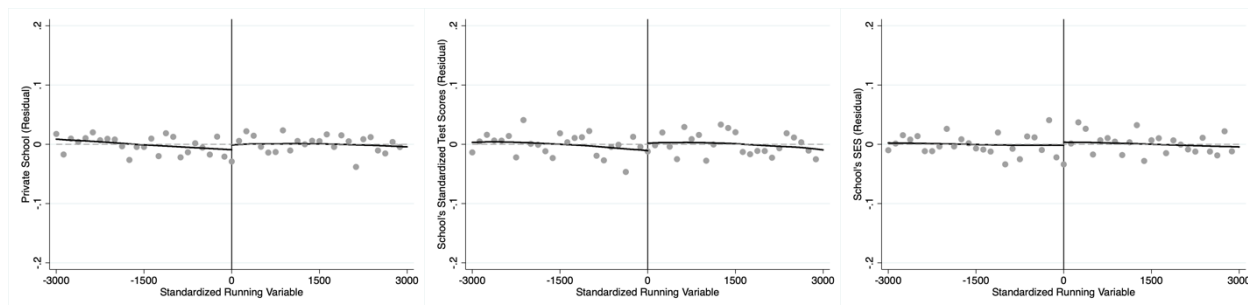
I further complement this formal conditional independence assumption testing with a graphical tool that looks at the relationship between outcome residuals—after regressing outcomes on the conditional vector—and the running variable. If the conditional independence assumption is correct, then the relationship between outcome residuals and the running variable should be flat, except possibly for a jump at the cutoff. Results can be found in Figure A.4. Black lines depict a fourth-order polynomial fit for control and treatment units separately, and grey dots represent the sample average for each disjoint bin. Consistent with the results reported in Table A.9, the relationship between outcome residuals and the running variable is essentially flat.

Finally, Figure A.5 plots the histogram of propensity score fitted values for treated and control observations above and below a common horizontal axis. The figure shows that there is substantial overlap to implement a propensity score estimate approach.

Table A.9:
Conditional Independence Test

	(1) School Private	(2) School Test Scores	(3) School SES	(4) School Add-On	(5) School Class Size	(6) School Distance (Miles)	(7) Student Language Second Grade	(8) Student Language Fourth Grade	(9) Student Math Fourth Grade
Panel A: Below the Cutoff Without Controls									
Eligible	-0.001 (0.004)	0.014*** (0.005)	0.024*** (0.005)	0.739*** (0.246)	0.016 (0.077)	0.018 (0.020)	0.013* (0.007)	0.004 (0.007)	0.018*** (0.007)
Mean Control	0.554	0.563	0.815	19.857	30.696	1.409	-0.081	0.156	0.187
Observations	24,973	24,211	24,180	24,860	24,618	10,059	23,211	21,352	21,417
Panel B: Below the Cutoff With Controls									
Eligible	-0.006** (0.003)	-0.005 (0.005)	-0.001 (0.004)	-0.083 (0.215)	-0.138** (0.065)	0.033* (0.019)	-0.005 (0.007)	-0.011 (0.007)	0.000 (0.006)
Mean Control	0.554	0.563	0.815	19.857	30.696	1.409	-0.081	0.156	0.187
Observations	24,973	24,211	24,180	24,860	24,618	10,059	23,211	21,352	21,417
Panel C: Above the Cutoff Without Controls									
Eligible	0.013*** (0.004)	0.014*** (0.005)	0.036*** (0.005)	1.307*** (0.266)	0.306*** (0.076)	-0.026 (0.021)	0.025*** (0.007)	0.025*** (0.007)	0.014** (0.007)
Mean Control	0.554	0.563	0.815	19.857	30.696	1.409	-0.081	0.156	0.187
Observations	24,973	24,211	24,180	24,860	24,618	10,059	23,211	21,352	21,417
Panel D: Above the Cutoff With Controls									
Eligible	-0.001 (0.003)	-0.004 (0.005)	-0.003 (0.004)	-0.268 (0.229)	-0.033 (0.066)	-0.047** (0.020)	0.007 (0.007)	0.012 (0.007)	-0.001 (0.006)
Mean Control	0.554	0.563	0.815	19.857	30.696	1.409	-0.081	0.156	0.187
Observations	24,973	24,211	24,180	24,860	24,618	10,059	23,211	21,352	21,417

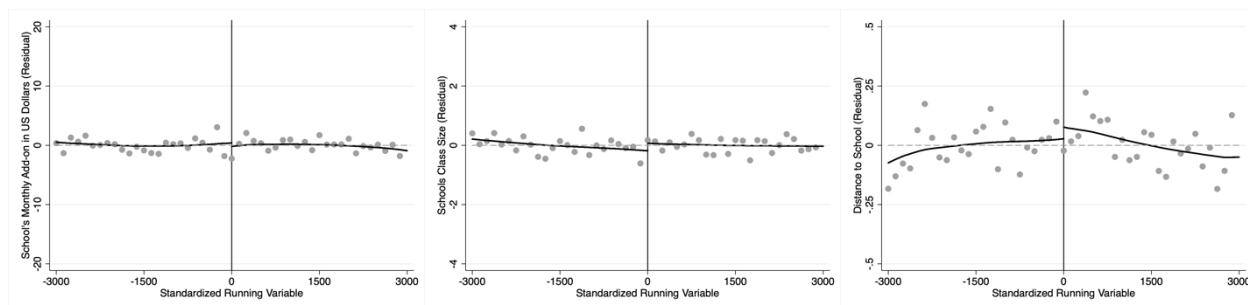
Notes: Table A.9 reports regression-based tests of the conditional independence assumption described in the text. Panels B and D show the coefficient on the father's education, household income, municipality of residence, and child's gender. Estimates use only observations to the left or right of the cutoff as indicated in column headings. Robust standard errors are reported in parentheses.



(A) Private School

(B) School's Test Scores

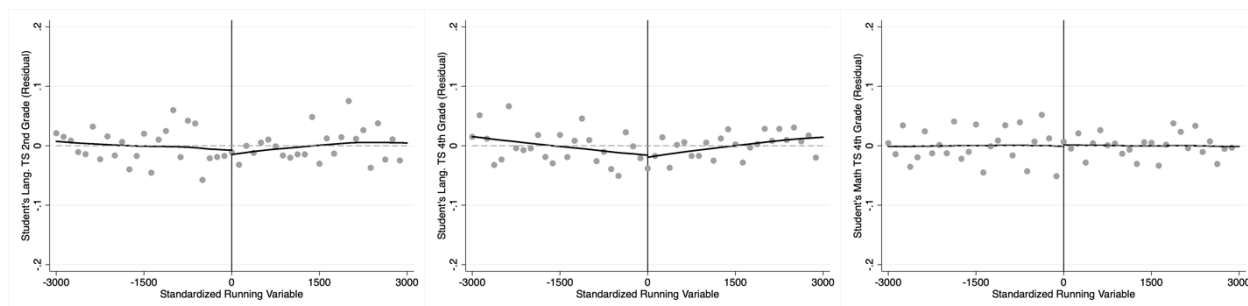
(C) School's SES



(D) School's Monthly Add-on

(E) School's Class Size

(F) Distance to School



(G) Student's Language Score
2nd Second Grade

(H) Student's Language Score
Fourth Grade

(I) Student's Math Score
Fourth Grade

Figure A.4: CIA Test

Notes: Figure A.4 looks at the relationship between outcome residuals—after regressing outcomes on the conditional vector—and the running variable. Grey dots present the average for individuals in equally spaced disjoint bins. Black lines depict a fourth-order polynomial fit for control and treatment units separately.

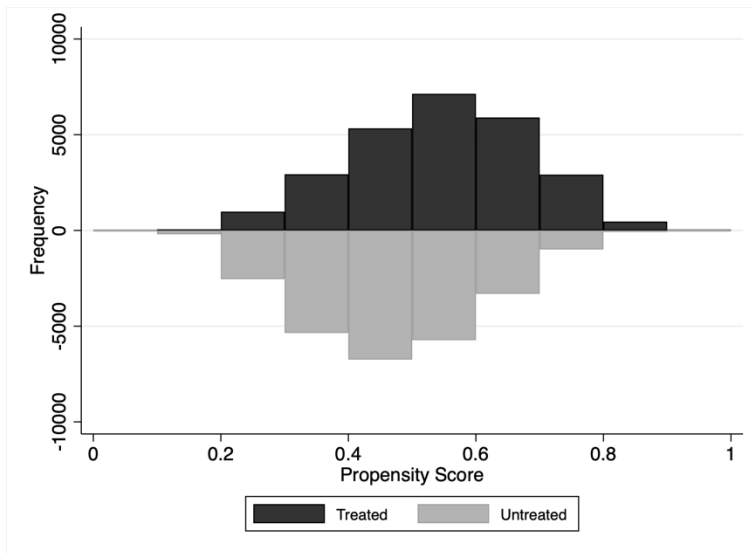


Figure A.5: Histogram of Estimated Propensity Score in the Window $[-3000,3000]$