Examining the Effects of Gifted Classes on Mathematics and Reading/Language Arts Achievement

Proposal

This paper examines whether increased gifted instruction in the content areas of math and Reading/Language Arts (RLA) increases student achievement for gifted students in a large southern urban school district. Meta-analyses of grouping students into gifted classes shows increases in the achievement of gifted students (Steenbergen-Hu et al., 2016). In addition, many researchers have found that increasing the rigor of academic courses increases student achievement. However, only a few studies have used experimental or quasi-experimental designs to test the effectiveness of gifted education (Author, 2011; 2015; Gavin et al., 2007; 2009). Although these studies present experimental evidence of positive effects of exposure to enriched academic instruction on gifted student achievement, a propensity score matching analysis of Early Childhood Longitudinal Study data found no evidence of a positive effect of gifted education for gifted students and no negative effect for non-gifted students (Adelson, 2012). Our study uses new evidence and a quasi-experimental design to deepen our understanding of the effects of gifted classes on achievement.

Methods

To evaluate the impact of gifted classes on the achievement of students, we compare the RLA and mathematics achievement of gifted students in three different settings: schools offering a fulltime gifted-only program with gifted classes in all subject areas, schools offering a part-time gifted-only program with gifted classes in mathematics, and schools offering a part-time gifted-only program with gifted classes in RLA. In part-time programs, students attend gifted classes in their school's focal content area and are in general education classes for other content areas. Therefore, we can compare the RLA and mathematics achievement of gifted students in gifted classes to the achievement of gifted students in general education classes.

We conducted six sets of analyses: a simple comparison of students' fifth-grade achievement in schools of each program type using a bivariate OLS, an OLS estimate with additional covariates, propensity score matching analyses with and without covariates, and multilevel matching with and without covariates.

The naïve OLS model served as a point of comparison; however, we expected the results of the standard OLS to be biased, due to selection issues. Students and schools with different program models might vary by socioeconomic status (SES), early academic achievement, English learner (EL) status, and demographic characteristics. To address this potential source of bias, our second set of analyses tested the effects of gifted instruction by adding a full set of covariates to the original OLS regression models.

However, OLS regression with covariates might not fully capture the differences between students and schools in different programs. It is possible that an OLS regression with covariates

could be biased if key confounding variables were not included in the OLS model. To address this issue, we employed propensity score weighting and a non-propensity score multilevel matching strategy.

In a third and fourth set of analyses, we used propensity score weighting with and without covariates to account for differences between the programs. In a fifth and sixth set of analyses, we used multilevel matching with and without covariates to control for differences between school and district. Our multilevel matching approaches address some recent critiques of propensity matching to better approximate a true experiment when using observational data (King & Nielson, 2019; Pimentel, Page, Lenard & Keele, 2018).

Data

This study used data from a large, ethnically, economically, and linguistically diverse urban school district in a state that mandates gifted education identification and programming. This district provided fulltime gifted programming or part-time gifted math or RLA instruction as a function of the school students attend, not student ability. The three types of gifted instructional delivery types in one district offered an ideal setting to examine the impact of gifted-only classes on identified gifted students.

We gathered administrative data for two cohorts of 4th graders in the district from 2015/16 and 2016/17. We gathered student data on gifted status, EL status, SES, and race/ethnicity, and 2nd-through-5th-grade achievement. Each cohort grade consists of over 20,000 students, with 16% gifted. Among 175 schools that had some gifted instruction, 47% of these schools provided fulltime instruction, 6% provided part-time math instruction, 17% provided part-time RLA instruction, and the remaining schools had gifted instruction taught by itinerant teachers or a mixed model that included some fulltime and some part-time instruction.

For both cohorts, we examined academic achievement scores in mathematics and RLA from 2nd through 5th grade, gifted student status, student poverty (measured by free and reduced-price lunch status (FRPL)), EL status, and student race/ethnicity for each student. We also examined school average socioeconomic status, percent gifted, percent EL, average school academic achievement, and the racial/ethnic composition of the school.

Results

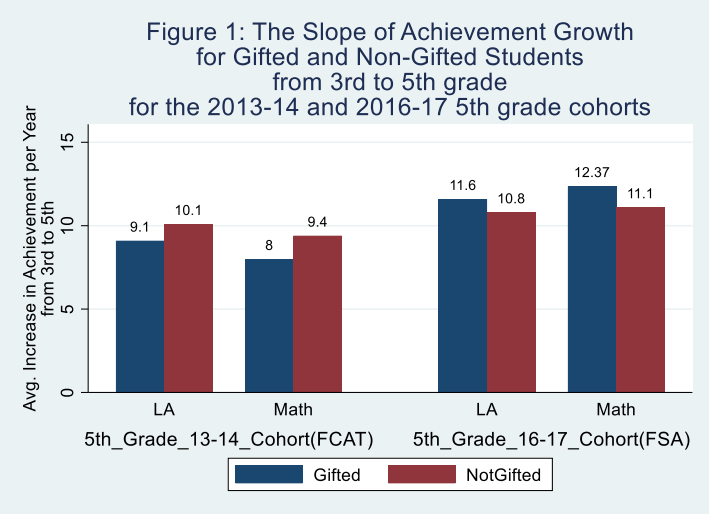
Our quantitative analysis revealed no discernable effects of gifted classes on either grade-level mathematics or RLA achievement (see Tables 3 and 4).

- There were no differences in RLA and math achievement across the various types of program offerings after controlling for the number of FRPL students in the school (see Tables 3 and 4).
- When we repeated these analyses with gain scores instead of 5th-grade achievement as the dependent variable, our results were similar (see Tables 3 and 4).

- There were no differences between gifted and general classes in math or RLA achievement in propensity score or multilevel matching models.
- Part-time RLA schools have lower levels of math achievement and higher levels of student poverty compared to fulltime and part-time math schools for all grades and cohorts (see Figures 2-5).
- Schools with part-time gifted math classes had lower levels of student poverty than full-time and part-time RLA schools. Our analysis of the descriptive characteristics between treatment and control groups found that there were large differences between schools and students in these two groups (see Tables 1-2). Our attempt to improve balance with weighted propensity score matching had only limited success. In contrast, multilevel matching provided a good balance between treatment and control groups with a standardized difference of .2 or less for most comparisons. However, this improved balance was at the cost of a dramatic reduction in sample size (see Tables 1-2).

Figures

Achievement Growth for Gifted vs. Non-Gifted Students in 2011/12 to 2013/14 vs. 2014/15 to 2016/17



Academic Achievement, School Poverty, and Student Poverty by Type of Gifted Program for 4th and 5th Grade Students in Two Cohorts

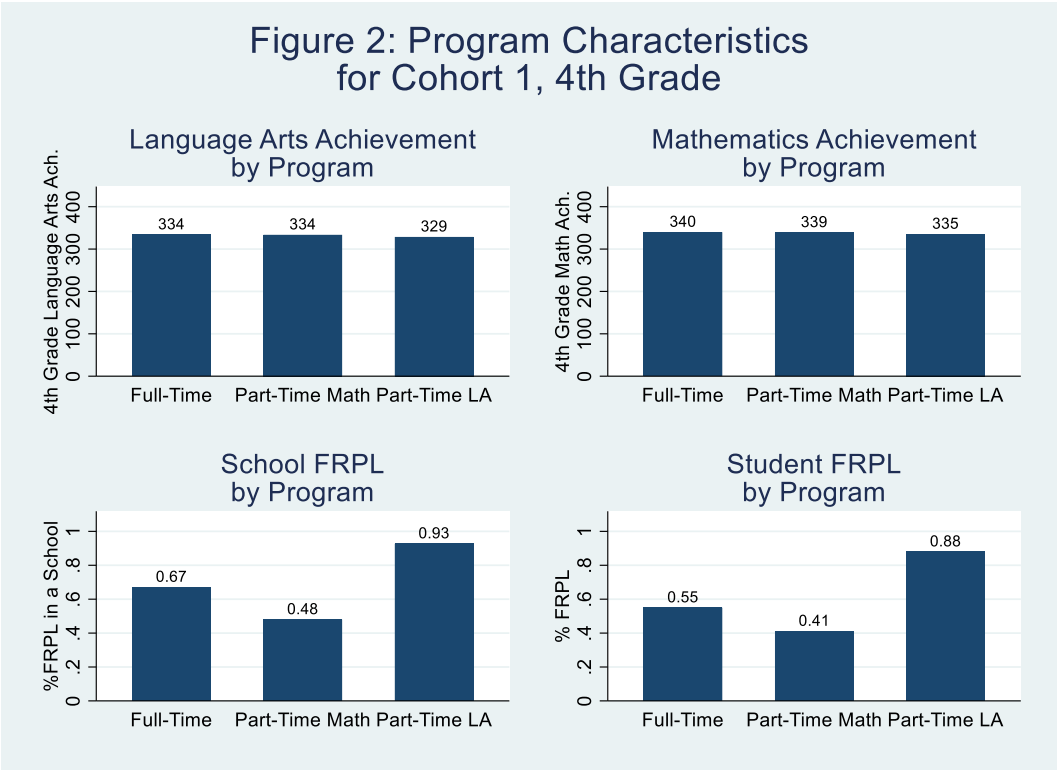


Figure 3: Program Characteristics for Cohort 1, 5th Grade

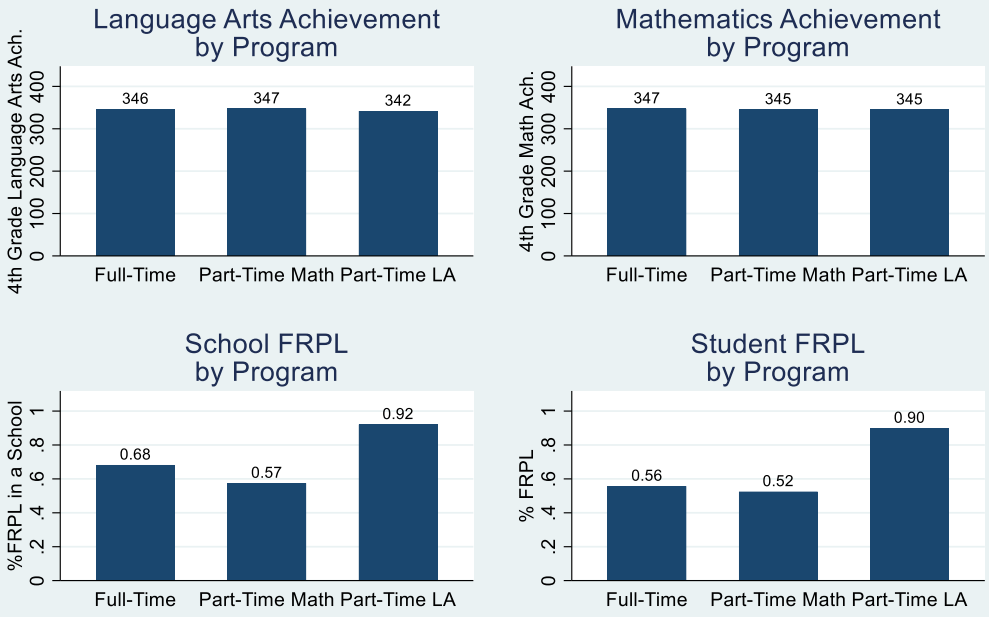


Figure 4: Program Characteristics for Cohort 2, 4th Grade

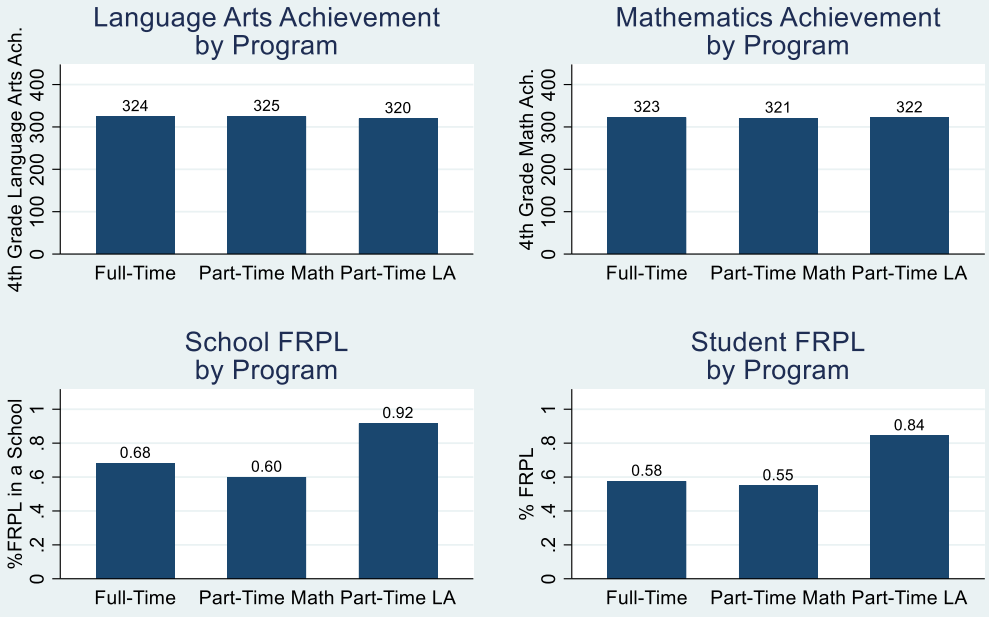
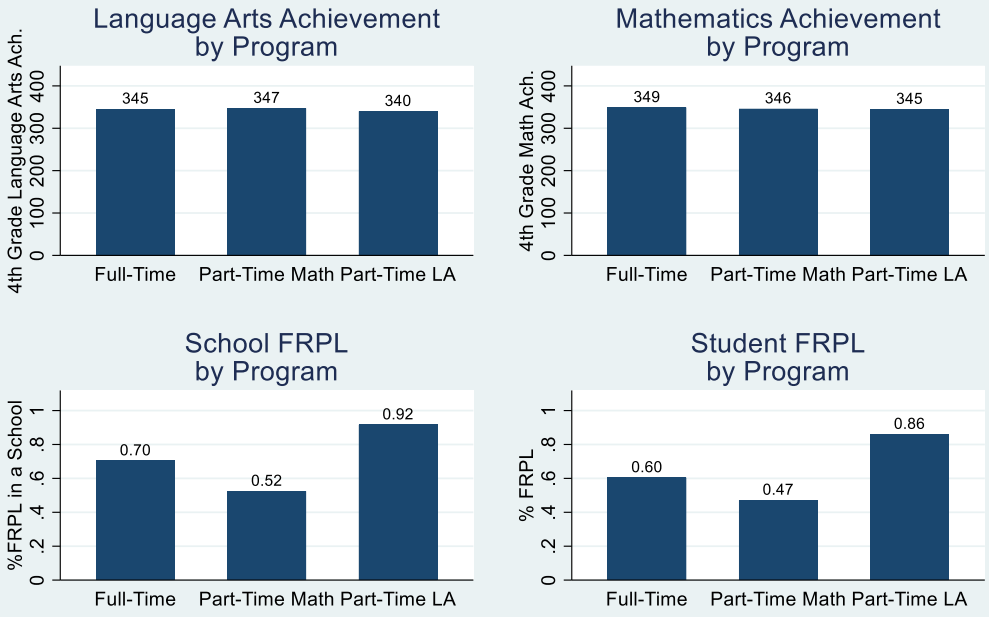


Figure 5: Program Characteristics for Cohort 2, 5th Grade



Tables

Table 1: *Mean Standardized Effect Size and Sample Size for Gifted Instruction in RLA Analyses for 4th and 5th Grades for 2 Cohorts*

Full-time vs. Part-time Math

Table 1
Mean Standardized Effect Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS	Propensity Score Weighting	Multilevel Matching
1	Grade 4 Cohort 1	0.31	0.35	0.16
2	Grade 5 Cohort 1	0.29	0.38	0.19
3	Grade 4 Cohort 2	0.25	0.31	0.33
4	Grade 5 Cohort 2	0.43	0.36	0.54

(Note: highlighted cells have adequate balance due to small differences between covariate means. We define an adequate balance as a mean standardized difference of less than .2)

Table 1B
Sample Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS		Propensity Score Weighting		Multilevel Matching	
		Students	Schools	Students	Schools	Students	Schools
1	Grade 4 Cohort 1	2247	92	2247	92	419	22
2	Grade 5 Cohort 1	2204	97	2204	97	424	30
3	Grade 4 Cohort 2	2076	97	2076	94	238	26
4	Grade 5 Cohort 2	2156	97	2156	94	119	14

Part-time RLA vs. Part-time Math

Table 1C
Mean Standardized Effect Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS	Propensity Score Weighting	Multilevel Matching
5	Grade 4 Cohort 1	0.63	0.41	0.29
6	Grade 5 Cohort 1	0.51	0.27	0.4
8	Grade 4 Cohort 2	0.52	0.26	0.25
9	Grade 5 Cohort 2	0.63	0.53	0.37

Table 1D
Sample Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS		Propensity Score Weighting		Multilevel Matching	
		Students	Schools	Students	Schools	Students	Schools
5	Grade 4 Cohort 1	569	41	569	41	92	14
6	Grade 5 Cohort 1	560	44	560	44	166	24
7	Grade 4 Cohort 2	577	44	577	45	192	26
8	Grade 5 Cohort 2	482	44	482	35	122	14

Table 2
Mean Standardized Effect Size and Sample Size for Gifted Instruction in Math Analyses for 4th and 5th Grades for 2 Cohorts

Full-time vs. Part-time RLA

Table 2A
Mean Standardized Effect Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS	Propensity Score Weighting	Multilevel Matching
9	Grade 4 Cohort 1	0.51	0.37	0.17
10	Grade 5 Cohort 1	0.5	0.32	0.12
11	Grade 4 Cohort 2	0.49	0.36	0.18
12	Grade 5 Cohort 2	0.55	0.21	0.15

(Note: highlighted cells have an adequate balance due to small differences between covariate means. We define an adequate balance as a mean standardized difference of less than .2)

Table 2B:
Sample Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS		Propensity Score Weighting		Multilevel Matching	
		Students	Schools	Students	Schools	Students	Schools
9	Grade 4 Cohort 1	2303	111	2303	111	512	56
10	Grade 5 Cohort 1	2291	111	2291	111	468	52
11	Grade 4 Cohort 2	2131	111	2131	111	500	56
12	Grade 5 Cohort 2	2252	111	2252	113	451	52

Part-time Math vs. Part-time RLA

Table 2C

Mean Standardized Effect Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS	Propensity Score Weighting	Multilevel Matching
13	Grade 4 Cohort 1	0.63	0.41	0.33
14	Grade 5 Cohort 1	0.5	0.24	0.33
15	Grade 4 Cohort 2	0.52	0.26	0.2
16	Grade 5 Cohort 2	0.63	0.53	0.27

(Note: highlighted cells have an adequate balance due to small differences between covariate means. We define an adequate balance as a mean standardized difference of less than .2).

Table 2D

Sample Size for Baseline, Propensity Score Weighting, and Multilevel Matching

#		OLS		Propensity Score Weighting		Multilevel Matching	
		Students	Schools	Students	Schools	Students	Schools
13	Grade 4 Cohort 1	566	41	566	41	107	14
14	Grade 5 Cohort 1	578	44	578	44	218	28
15	Grade 4 Cohort 2	576	44	576	45	183	26
16	Grade 5 Cohort 2	482	44	482	35	127	14

Table 3

Effect of Gifted Instruction in RLA for 4th and 5th Grades for 2 Cohorts

#	Comparison	Cohort	Grade	OLS				Propensity Score Weighting				Multilevel Matching			
				Time 2		DID(Gain)		Time 2		DID(Gain)		Time 2		DID(Gain)	
				No Covar	Covar	No Covar	Covar	No Covar	Covar	No Covar	Covar	No Covar	Covar	No Covar	Covar
1	**F vs. PM** Gifted RLA and Math Programs (Full Time) vs. No Gifted RLA Program	1	4th	0.2	0.4	0.49	1.31	3.89	0.77	1.15	0.66	2.07	2.83	2.56	3.64
				[0.92]	[0.83]	[0.76]	[0.82]	[1.53]	[1.06]	[1.14]	[1.08]	[1.79]	[1.99]	[1.89]	[2.08]
2		1	5th	-0.15	-2.3	-1.57	-2.03	4.45	0.58	-0.56	-0.47	-0.52	0.04	0.43	-0.28
				[1.03]	[0.97]	[0.77]	[0.87]	[1.98]	[1.4]	[1.27]	[1.05]	[1.88]	[1.57]	[1.79]	[1.48]
3		2	4th	1.12	0.77	1.7	1.86	1.87	1.82	-0.45	-1.54	-2.49	-3.83	-1.2	-2.64
				[0.95]	[0.89]	[0.8]	[0.89]	[1.52]	[1.16]	[1.31]	[1.25]	[1.54]	[1.68]	[1.5]	[2.09]
4		2	5th	-1.7	-2.46	-2.94 *	-3.69 *	-2.97	-1.13	-4.66	-3.77	-0.68	-5.13	-4.61	-0.83
				[1.1]	[1.13]	[0.91]	[1.09]	[4.84]	[3.68]	[3.79]	[3.5]	[3.59]	[7.96]	[3.02]	[6.01]
5	**PM vs. PRLA** Gifted RLA Program Only vs. No Gifted RLA Program	1	4th	-4.96 *	-0.2	1.59	4.16 *	-3.88	-0.38	2.73	4.07 *	-1.87	4.31	2.37	7.2
				[1.23]	[1.5]	[0.91]	[1.37]	[1.58]	[1.5]	[1.14]	[1.43]	[3.06]	[3.29]	[2.89]	[3.14]
6		1	5th	-4.38 *	-0.26	-0.16	0.98	-1.65	0.1	0.87	1.67	1.97	-0.41	4.56 *	3.45
				[1.35]	[1.44]	[1.01]	[1.37]	[1.69]	[1.54]	[1.11]	[1.22]	[2.45]	[3.49]	[1.72]	[3.25]
7		2	4th	-3.82 *	0.1	1.44	-0.41	-1.53	-0.36	1.54	0.69	0.28	1.09	2.85	4.55
				[1.2]	[1.28]	[1.02]	[1.32]	[1.43]	[1.32]	[1.54]	[1.55]	[2.37]	[2.17]	[2.16]	[2.32]
8		2	5th	-7.11 *	-5.86 *	-3.07 *	-5.27 *	-8.51 *	-7.65 *	-3.81 *	-6.25 *	-5.83	-4.39	-4.25	-7.95
				[1.44]	[2.05]	[1.1]	[1.91]	[1.63]	[2.11]	[1.4]	[2.18]	[3.43]	[6.66]	[2.68]	[6.07]

Notes:

- F vs. PM = Full-time Instruction vs. Part-time Math Instruction; PM vs. PRLA = Part-time Math instruction vs. Part-time RLA Instruction;
- *= *p*-value<=.05; Statistically significant covariates are highlighted; Standard Error in Brackets;
- 64 comparisons; Only treatment coefficients shown; full set of covariates = FRPL Status, EL Status, Black, Hispanic, Asian, and other, 2nd grade math and RLA achievement at the student level and % gifted, % EL, % FRPL, % Black, % Hispanic, average 2nd grade math and RLA scores at the school level, Multilevel Matching based on the R package multi-match, which consists of a two-staged multilevel matching procedure (Pimentel, Page, & Keele, 2018); propensity score weight based on a logistic regression of all covariates; For the propensity score matching the extreme values were trimmed and stabilized based on the methods in Harder, Stuart, & Anthony (2010).

Table 4

Effect of Gifted Instruction in Math for 4th and 5th Grades for 2 Cohorts

#	Comparison	Cohort	Grade	OLS				Propensity Score Weighting				Multilevel Matching			
				Time 2		DID(Gain)		Time 2		DID(Gain)		Time 2		DID(Gain)	
				No Covar	Covar	No Covar	Covar	No Covar	Covar	No Covar	Covar	No Covar	Covar	No Covar	Covar
9	**F vs. PRLA** Gifted Math and Gifted RLA Programs vs. No Gifted Math Program	1	4th	4.69 *	0.29	0.85	-0.45	2.83	-0.1	-0.01	-1.51	-1.35	-2.16	-3.46	-4.51
				[1.06]	[1.04]	[0.82]	[0.95]	[1.67]	[1.3]	[0.94]	[0.97]	[2.17]	[2.28]	[1.94]	[1.98]
10		1	5th	2.36	-2.45	-2.41 *	-1.57	-2.02	-3.37 *	0.53	0.42	-2.71	-4 *	-1.89	-2.6
				[1.1]	[1.12]	[0.8]	[0.95]	[1.51]	[1.28]	[1.51]	[1.49]	[1.55]	[1.53]	[1.86]	[1.81]
11		2	4th	3.17 *	-1.95	1.93	-1.03	-0.43	-1.96	0.21	-0.88	-2.93	-3.18	-0.85	-1.33
				[0.96]	[0.95]	[0.87]	[1]	[1.3]	[1.23]	[1.13]	[1.22]	[1.86]	[2.06]	[1.26]	[1.45]
12		2	5th	4.11 *	-1.33	3.04 *	0.59	-0.18	0.52	2.94	2.79	1.27	0.4	2.44	0.83
				[1.1]	[1.1]	[0.9]	[1.03]	[2.79]	[2.18]	[1.76]	[1.64]	[2.22]	[2.1]	[1.93]	[1.96]
13	**PM vs. PRLA** Gifted Math Program Only vs. No Gifted Math Program	1	4th	3.87	-4.51	1.14	-3.93	2.15	-3.77	-0.23	-3.65	-4.73	-5.71	-4.48	-2.53
				[1.52]	[1.85]	[1.15]	[1.69]	[1.91]	[1.7]	[1.4]	[1.54]	[4.87]	[6.71]	[2.58]	[6.06]
14		1	5th	0.48	-1.44	-3.55 *	0.9	-1.73	-2.88	0.06	1.47	0.08	1.35	-2.25	0.05
				[1.46]	[1.76]	[1.12]	[1.52]	[2.25]	[2.13]	[1.62]	[1.67]	[3.38]	[2.89]	[2.79]	[2.61]
15		2	4th	1.63	1.29	1.97	1.95	1.1	1.22	1.57	1.79	3.35	5.46	5.47	5.78
				[1.34]	[1.49]	[1.16]	[1.51]	[1.77]	[1.47]	[1.46]	[1.5]	[3.28]	[3.03]	[3.2]	[3.49]
16		2	5th	1.17	3.31	-3.48 *	0.18	1.57	2.11	-3.36	-1.12	3.03	1.7	-4.92	-4.32
				[1.6]	[2.34]	[1.1]	[1.89]	[1.83]	[2.62]	[1.65]	[2.39]	[2.89]	[3.99]	[3.23]	[5.84]

Notes:

- F vs. PRLA = Full-time Instruction vs. Part-time RLA Instruction; PM vs. PRLA = Part-time Math instruction vs. Part-time RLA instruction;
- *= *p*-value<=.05; Statistically significant covariates are highlighted; Standard Error in Brackets;
- 64 comparisons; Only treatment coefficients shown; full set of covariates = Free or reduced-price lunch Status (FRPL), English learner Status (EL), Black, Hispanic, Asian, and other, 2nd grade math and RLA achievement at the student level and % gifted, % EL, % FRPL, % Black, % Hispanic, average 2nd grade math and RLA scores at the school level, Multilevel Matching based on the R package multi-match, which consists of a two-staged multilevel matching procedure (Pimentel, Page, & Keele, 2018); propensity score weight based on a logistic regression of all covariates; For the propensity score matching the extreme values were trimmed and stabilized based on the methods in Harder, Stuart, & Anthony (2010).

Descriptive Statistics by Program Type

Table 5
Cohort 1 – 5th grade in 2016/2017

Variables	Fulltime		Part-time Math		Part-time RLA	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Student Level						
% FRPL	0.56	0.50	0.52	0.50	0.90	0.30
% Black	0.08	0.26	0.29	0.45	0.30	0.46
% Hispanic	0.72	0.45	0.51	0.50	0.66	0.48
% Asian	0.02	0.15	0.04	0.19	0.00	0.06
% Other	0.01	0.11	0.02	0.14	0.01	0.08
% White	0.17	0.38	0.15	0.36	0.03	0.18
2nd grade LA	657.13	33.27	651.50	33.43	648.20	32.38
RLA Ach. 14/15	323.00	14.10	321.45	14.04	317.51	13.95
RLA Ach. 15/16	333.72	13.91	333.10	14.33	328.08	14.38
RLA Ach. 16/17	346.12	15.42	347.10	15.92	341.73	15.59
2nd grade Math	649.47	35.05	650.62	34.31	640.59	34.75
Math Ach. 14/15	321.71	15.27	319.86	14.02	319.86	17.22
Math Ach. 15/16	339.69	17.14	338.93	17.88	334.92	18.19
Math Ach. 16/17	347.16	18.50	345.29	17.42	344.74	17.22
School Level						
% FRPL in school	0.68	0.20	0.57	0.38	0.92	0.06
% Gifted in school	0.27	0.13	0.34	0.25	0.13	0.04
% Black in school	0.10	0.17	0.29	0.36	0.28	0.34
% Hispanic in school	0.75	0.23	0.52	0.30	0.69	0.34
% Asian in school	0.02	0.02	0.02	0.02	0.01	0.01
% Other in school	0.01	0.01	0.01	0.03	0.01	0.01
% White in school	0.12	0.14	0.15	0.15	0.02	0.02
2nd grade RLA Ach., School Avg.	624.70	12.03	625.94	25.10	607.71	8.49
2nd grade Math Ach., School Avg.	617.29	14.00	620.15	27.09	600.60	13.77

Table 6
Cohort 2 – 5th grade in 2017/2018

Variables	Fulltime		Part-time Math		Part-time RLA	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Student Level						
% FRPL	0.60	0.49	0.47	0.50	0.86	0.35
% Black	0.06	0.24	0.31	0.46	0.31	0.46
% Hispanic	0.75	0.43	0.44	0.50	0.65	0.48
% Asian	0.02	0.15	0.04	0.19	0.00	0.06
% Other	0.01	0.11	0.01	0.10	0.01	0.12
% White	0.15	0.36	0.21	0.41	0.02	0.15
2nd grade RLA	657.55	32.79	656.72	33.39	648.34	34.32
RLA Ach. 15/16	324.17	14.33	325.23	15.21	319.66	15.94
RLA Ach. 16/17	334.84	14.22	333.60	14.12	329.57	15.08
RLA Ach. 17/18	345.35	14.41	347.06	15.68	339.85	15.32
2nd grade Math	650.26	34.92	649.65	33.43	637.05	37.43
Math Ach. 15/16	322.74	15.79	319.53	14.32	320.93	15.88
Math Ach. 16/17	339.96	15.44	339.61	13.94	335.05	16.89
Math Ach. 17/18	348.95	17.48	346.04	15.10	344.76	18.40
School Level						
% FRPL in school	0.70	0.22	0.52	0.38	0.92	0.09
% Gifted in school	0.32	0.16	0.41	0.18	0.21	0.09
% Black in school	0.08	0.15	0.33	0.40	0.35	0.39
% Hispanic in school	0.78	0.22	0.44	0.25	0.63	0.38
% Asian in school	0.02	0.02	0.02	0.01	0.00	0.01
% Other in school	0.01	0.01	0.01	0.01	0.01	0.01
% White in school	0.12	0.14	0.20	0.16	0.02	0.03
2nd grade RLA Ach., School Avg.	625.82	12.45	631.54	19.72	612.38	9.49
2nd grade Math Ach., School Avg.	618.69	13.40	625.82	24.81	603.33	12.10

References

Adelson, J. L. (2012). Examining relationships and effects in gifted education research: An introduction to structural equation modeling. *Gifted Child Quarterly*, *56*, 47-55. doi.org/10.1177/0016986211424132

Gavin, M. K., Casa, T. M., Adelson, J. L., Carroll, S. R., & Sheffield, L. J. (2009). The impact of advanced curriculum on the achievement of mathematically promising elementary students. *Gifted Child Quarterly, 53,* 188-202. doi. 10.1177/0016986209334964

Gavin, M. K., Casa, T. M., Adelson, J. L., Carroll, S. R., Sheffield, L. J., & Spinelli, A. M. (2007). Project M^3: Mentoring mathematical minds—A research-based curriculum for talented elementary students. *Journal of Advanced Academics, 18,* 566-585.

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods*, *15*(3), 234-249. doi: 10.1037/a0019623

King, G., & Nielsen, R. (2019). Why propensity scores should not be used for matching. *Political Analysis*, 1–20. doi.org/10.1017/pan.2019.11

Kulik, J. A. (2003). Grouping and tracking. In N. Colangelo & G. A. Davis (Eds). *Handbook of gifted education* (3rd ed., pp. 268-281). Boston, MA: Allyn & Bacon.

Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). New Haven, CT: Yale University Press.

Pimentel, S. D., Page, L. C., Lenard, M., & Keele, L. (2018). Optimal multilevel matching using network flows: An application to a summer reading intervention. *The Annals of Applied Statistics*, *12*, 1479-1505. doi.org/10.1214/17-AOAS1118

Slavin, R. E., & Karweit, N. L. (1985). Effects of whole class, ability grouped, and individualized instruction on mathematics achievement. *American Educational Research Journal*. doi.org/10.3102/00028312022003351

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research*, *86*, 849-899. doi/abs/10.3102/0034654316675417