The Research Alliance for New York City Schools

Assessing the Early Impact of School of One: Evidence from Three School-Wide Pilots
Technical Appendices

Rachel Cole
New York University

James J. Kemple
The Research Alliance for New York City Schools

Micha D. Segeritz
The Research Alliance for New York City Schools

June 2012

CONTENTS

APPENDIX A

COMPARATIVE INTERRUPTED TIME SERIES RESEARCH DESIGN

Overview

The impact analysis in this report is based on a design known as a comparative interrupted time series, a method used widely in education research and evaluation to assess the impact of school-wide programs and systemic policies on student outcomes.[1] The central strength of this methodology is that it accounts for many factors that may have produced changes in math achievement in the So1 schools instead of or in addition to the implementation of the school-wide So1 program in the 2010-2011 school year. The goal of accounting for these factors is to construct the best estimate of math achievement levels that were likely to have occurred in the So1 schools in the absence of the program. This alternative is known as a counterfactual. The analyses conducted for this chapter are based on a particularly strong counterfactual in that it accounts for many important alternative influences on student test scores that may have been present over and above the implementation of So1 in the 2010-2011 school year. A strong counterfactual increases confidence that the findings from the analyses constitute rigorous evidence of effects, or lack of effects, from the program.

There are several potential influences on math test scores that must be controlled for by the comparative interrupted time series analysis:

- Math curricula and teaching strategies that were underway prior to SO1 and may have helped improve or depress students' math achievement. Math test scores in the SO1 schools (and across New York City) started improving even before SO1 was implemented. Thus, it is likely that these trends would have continued even if SO1 had never been developed or introduced into these schools. The interrupted time series analysis isolates changes in test score trends that occurred in the SO1 schools in the 2010-11 school year over and above what would have occurred had the prior trends continued.
- Citywide and state reforms were aimed at improving math achievement across New York City. It is possible that the accountability mandates and school improvement initiatives required under Children First reforms beginning in 2002 produced improvements in student test scores independently of the reforms such as SO1. Similarly, there may be other federal or state policies aimed at school improvement that may cause test score improvements to continue into the 2010-11 school year and beyond. The comparative interrupted time series analysis isolates changes in test scores that occurred in SO1 during this period over and above those that occurred in other similar schools in New York City that were subject to the same policies, mandates, and reform initiatives.
- Changes in the state test, scoring methods, or performance criteria and increasing familiarity with the assessments and their frameworks. It is possible that the state

[1] This statistical methodology has been used widely in education research and evaluation (see Bloom, 1999 and Shadish, Cook, & Campbell, 2002). As in this paper, comparative interrupted time series analyses have been applied primarily to study broad systemic policies and interventions such as the federal No Child Left Behind Act of 2002 (see Dee & Jacob, 2008 and Wong, Cook, & Steiner, 2009), accountability systems (see Jacob, 2005) and comprehensive school reforms such as Accelerated Schools (see Bloom, 2001) and Talent Development High Schools (see Kemple, Herlihy, & Smith, 2005).

assessments in math became easier over time or that teachers and students became increasingly familiar with test content, scoring methods, and performance criteria. By comparing test score trends in SO1 schools with those of other similar schools in New York City (schools that used the same math tests for students in grades 6 through 8 and were subject to the same scoring methods and standards over time), the method can hold constant the independent effect of changes in the test or scoring criteria.

- Changes in the composition of the schools may have impacted math achievement. It is possible that in the 2010-2011 school year the particular students attending the SO1 schools were different in substantive ways form those that typically attended the schools in the past. For example, if the 2010-2011 students had lower prior achievement than that of students typically attending the schools, we might expect lower test performance in 2011 than would be predicted by the test score trend.

In short, the counterfactual for this analysis is the estimated test scores for SO1 schools in the 2010-11 school year controlling for: 1) the continuation of test score trends underway in New York City schools prior to that year; 2) the deviation of 2011 test scores from the baseline trends in similar schools; and 3) other measured differences in school characteristics between SO1 schools and a matched comparison group of New York City middle schools. This counterfactual represents the best available estimate of test score trends that were likely to have occurred for SO1 schools in the absence of the program. Thus, the best evidence of effects from these reforms is derived from the difference between the test score trends that actually occurred in SO1 schools and the estimated counterfactual trends.

Comparative Interrupted Time Series Analysis

Our analysis estimates a linear baseline comparative interrupted time-series model:

$$Y_{ijt} = INT + \beta_1 * YR_{ijt} + \beta_2 * 2011_{ijt} + \beta_3 * X_{ijt} + \beta_4 * So1_{ijt} + \beta_5 * So1_{ijt} * YR_{ijt} + \beta_6 * So1_{ijt} * 2011_{ijt} + \epsilon_j + \epsilon_t \varepsilon_{ijt}$$

Where:

Y_{ijt}	=	Test score for student *i* in school *j* in year *t.*
YR_{ijy}	=	Year of observation for student *i* in school *j*, where -4, -3, -2, -1, and 0 correspond to 2006 - 2010, respectively and 1 corresponds to 2011.
$Post2010_{ijt}$	=	1 if observation for student *i* in school *j* is from 2011, 0 if observation is from 2006-2010.
X_{ijt}	=	Vector of predictors of individual student characteristics for student *i* in school *j* in year *t.*
$So1_{ijt}$	=	1 if school *j* is a SO1 school, 0 otherwise.
$So1_{ijt} * YR_{ijt}$	=	Year of observation for SO1 school *j*, where -4, -3, -2, -1, and 0 correspond to 2006 - 2010, respectively and 1 corresponds to 2011.
$So1_{ijt} * Post2010_{ijt}$	=	1 if school *j* is a SO1 school and the year is 2011, 0 otherwise.
ϵ_j	=	error associated with the school random effect.

ϵ_t = error associated with the year random effect.

ε_{ijt} = random error for student *i* in school *j* in year *t.*

In addition to the individual level random variation $e_{i(tj)}$, the model takes two additional sources of random variation into account: school level random variation, ϵ_j, and random variation across years, ϵ_t. The random variation across years, u_t, accounts for variation in test score levels between years due to fluctuations in test difficulty. Similarly, the school level random effect, u_j, takes into account that students within schools may be more similar to each other than between schools. The two additional random effects account for the clustering of random errors within years and within schools and guarantee to correct standard error estimation of the model.

The comparative interrupted times series analysis proceeds in three stages.[2] The following discussion details these stages and explains the intuition underlying this methodology.

Stage 1: Math Test Score Trends for SO1 Schools

The first stage in the impact analysis compares math test scores in the 2010-11 school year in SO1 schools with a continuation of the math test score trend from 2006 through 2010 in the same schools. The difference between the observed score in 2011 and the estimated score for that year provides an initial indication of a change in math achievement concurrent with the implementation of SO1. However, we cannot necessarily attribute this deviation to SO1 since other district-wide reforms and policies may have come on line during this period and influenced math achievement independent of SO1.

Stage 2: Comparison Schools and their Math Test Score Trends

The second stage of the analysis begins with the identification of matched comparison schools with characteristics and test-score trajectories for 6th, 7th, and 8th grade prior to the 2010-2011 school year that are similar to those of each SO1 school.[3] Descriptive statistics comparing all three SO1 schools and the comparison schools can be found below in Table A1. Tables comparing each SO1 school with its comparison schools can be found below in Tables A2-A4. The comparison schools selected are the six schools that are most similar to each SO1 school on an index constructed of test-score trajectories and characteristics.[4] The matching process prioritized finding schools with similar average test scores and test score trends, but also

[2] This section draws on Howard Bloom's methodological work (1995, 1999, 2003). In the body of the report, the description of the research design is mostly conceptual and non-technical. In the footnotes and appendix we provide additional methodological and technical information.

[3] Potential comparison schools included only the 189 schools with a middle school grade configuration (Grades 6-8) that operated continuously between 2006 and 2011.

[4] Based on the concept of "Euclidian distances" used in many cluster analyses, the similarity index captures the multi-dimensional differences between each SO1 school and each potential comparison school based on important background and performance characteristics. The index was constructed by weighting the previous test scores trends and test score levels each by one third, and the combined demographic characteristics by one third. We then selected the six best matching schools as comparison schools for each SO1 school. Six comparison schools were chosen for each of the three SO1 schools to decrease the impact of potential idiosyncratic test score trends of one of the comparison schools.

included some attention to demographics. Then we compare math test scores in the 2010-11 school years in each set of comparison schools with a continuation of the math test score trend from 2006 through 2010 in the same schools. Since these schools were not exposed to SO1, deviations from the baseline trend would be due to other reforms or initiatives being implement across the district or in selected middle schools like these.

Once we selected the comparison schools for each SO1 school we excluded groups of students that were not exposed to SO1 from our analysis: students receiving bilingual instruction and students with special education needs. For the comparison schools matched to School C, we included only Grade 6 in our analysis.

Table A1
Characteristics of Students Served in SO1 Schools and Comparison Schools
(2009-2010 School Year)

Characteristic	SO1 Schools	Comparison Schools
ELL status[a]		
Bilingual (%)	9	2
English as a Second Language (%)	14	13
Non-ELL (%)	77	85
Special education status		
Small class (%)	5	6
Mainstream (%)	11	10
General education (%)	84	84
Race/ethnicity		
Black (%)	17	16
Hispanic (%)	32	43
Asian (%)	41	26
White (%)	10	14
Gender		
Female (%)	45	49
Male (%)	55	51
Grade		
6th Grade (%)	29	30
7th Grade (%)	35	33
8th Grade (%)	37	36
Over age for grade[b] (%)	31	23
Free/reduced lunch (%)	82	54
2010 attendance rate (%)	93	93
Average scaled score on 2010 NYS math test	670	675
Average yearly trend in math scores[c], 2006-2010	7	7
Total enrollment	2,498	19,490

Source: Research Alliance analysis of student characteristics.
[a] ELL designates English Language Learners.
[b] Over age for grade designates students who have repeated grades.
[c] The average yearly trend in math scores gives the typical yearly change in the school's average test scores in the five years before SO1 was implemented.

Table A1 above shows the similarities between the three SO1 schools to the eighteen comparison schools, particularly in their average test scores and their test score trends. At the same time, there are some modest differences in demographic characteristics, most notably, the percentage of students who were Hispanic and Asian, the percentage of students eligible for free and reduced lunch, and the percentage of students receiving bilingual instruction. This reflects how similarities in test scores and trends were prioritized over demographic similarities in our matching process. Tables A2 through A4 show this comparison for each of the three SO1 schools and their comparison schools.

Table A2
Characteristics of Students Served in School A and its Comparison Schools
(2009-2010 School Year)

Characteristic	School A	Comparison Schools
ELL status[a]		
Bilingual (%)	17	1
English as a Second Language (%)	18	18
Non-ELL (%)	65	81
Special education status		
Small class (%)	3	5
Mainstream (%)	14	9
General education (%)	83	86
Race/ethnicity		
Black (%)	5	8
Hispanic (%)	12	38
Asian (%)	83	42
White (%)	1	11
Gender		
Female (%)	43	49
Male (%)	57	51
Grade		
6th Grade (%)	28	28
7th Grade (%)	35	35
8th Grade (%)	37	37
Over age for grade[b] **(%)**	34	21
Free/reduced lunch (%)	85	57
2010 attendance rate (%)	96	94
Average scaled score on 2010 NYS math test	681	676
Average yearly trend in math scores[c]**, 2006-2010**	4	6
Total enrollment	856	8,125

Source: Research Alliance analysis of student characteristics.

[a] ELL designates English Language Learners.

[b] Over age for grade designates students who have repeated grades.

[c] The average yearly trend in math scores gives the typical yearly change in the school's average test scores in the five years before SO1 was implemented.

Table A3
Characteristics of Students Served in School B and its Comparison Schools
(2009-2010 School Year)

Characteristic	School B	Comparison Schools
ELL status[a]		
Bilingual (%)	2	0
English as a Second Language (%)	8	6
Non-ELL (%)	90	94
Special education status		
Small class (%)	4	5
Mainstream (%)	8	10
General education (%)	88	85
Race/ethnicity		
Black (%)	16	16
Hispanic (%)	22	41
Asian (%)	34	20
White (%)	28	22
Gender		
Female (%)	44	50
Male (%)	56	50
Grade		
6th Grade (%)	28	32
7th Grade (%)	36	33
8th Grade (%)	36	35
Over age for grade[b] **(%)**	18	17
Free/reduced lunch (%)	75	40
2010 attendance rate (%)	94	94
Average scaled score on 2010 NYS math test	678	680
Average yearly trend in math scores[c]**, 2006-2010**	5	5
Total enrollment	880	8,070

Source: Research Alliance analysis of student characteristics.

[a] ELL designates English Language Learners.

[b] Over age for grade designates students who have repeated grades.

[c] The average yearly trend in math scores gives the typical yearly change in the school's average test scores in the five years before SO1 was implemented.

Table A4
Characteristics of Students Served in School C and its Comparison Schools
(2009-2010 School Year)

Characteristic	School C	Comparison Schools
ELL status[a]		
Bilingual (%)	9	11
English as a Second Language (%)	16	15
Non-ELL (%)	75	74
Special education status		
Small class (%)	8	10
Mainstream (%)	12	10
General education (%)	80	80
Race/ethnicity		
Black (%)	32	34
Hispanic (%)	66	64
Asian (%)	1	1
White (%)	0	1
Gender		
Female (%)	48	47
Male (%)	52	53
Grade		
6th Grade (%)	30	32
7th Grade (%)	33	32
8th Grade (%)	37	36
Over age for grade[b] (%)	44	40
Free/reduced lunch (%)	85	81
2010 attendance rate (%)	89	90
Average scaled score on 2010 NYS math test	650	657
Average yearly trend in math scores[c], 2006-2010	11	10
Total enrollment	762	3295

Source: Research Alliance analysis of student characteristics.

[a] ELL designates English Language Learners.

[b] Over age for grade designates students who have repeated grades.

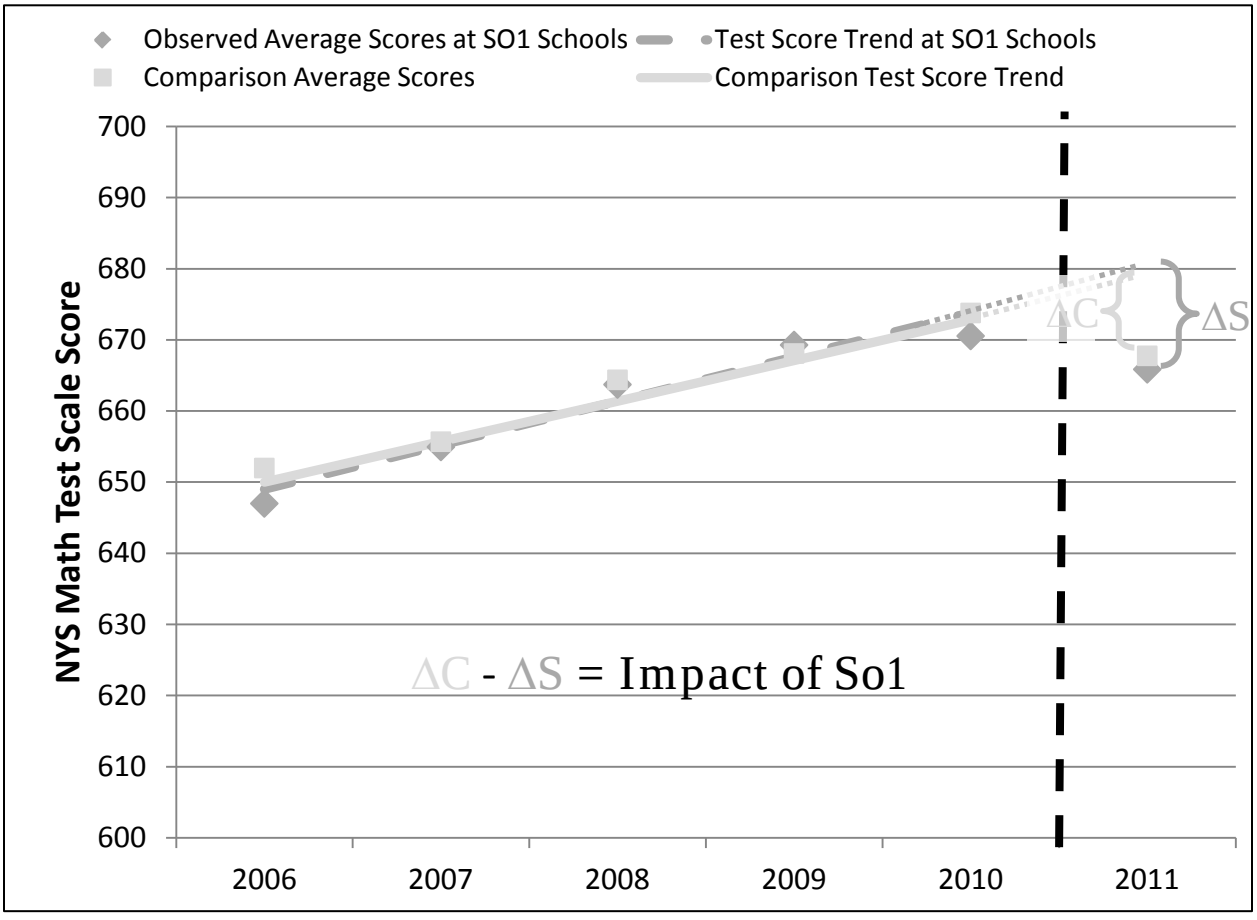
[c] The average yearly trend in math scores gives the typical yearly change in the school's average test scores in the five years before SO1 was implemented.

Stage 3: Comparing Changes in Test Score Trends for SO1 and Comparison Schools

In the final stage of the analysis, we compare the differences estimated in Stage 1 with the differences estimated in Stage 2. Thus, this so-called "difference-in-difference" approach contrasts the drops in 2011 for the SO1 schools to the drops for the comparison schools. This estimate represents the best indication of the impact SO1 has on student math test scores over and above the influence of prior initiatives and trends and simultaneous interventions that may be underway across the district or in schools like those being served by SO1.

Figure A1 shows a theoretical example of the the baseline and projected trends for 6th grade students at the SO1 schools and their matched comparison school counterparts. The figure illustrates a high degree of similarity both in the levels of student performance and in the year-to-year growth in test scores across both sets of schools from 2006 to 2010. This suggests that one may have a high degree of confidence that subsequent differences that emerged between the schools in 2011 are likely to be due to one school being exposed to SO1 and the comparison schools not being exposed. The figure illustrates the deviations between the 2011 scores and the projected trends, ΔC and ΔS, that will be compared to estimate the impact of SO1.

Figure A1
Comparative Interrupted Time Series
Theoretical Example



Source: Research Alliance theoretical example.

Pooling results across schools and grade levels

The analyses discussed above will result in separate impact estimates for 6th grade in each SO1 school. Alone, these estimates are likely to be unreliable indicators of SO1's overall effectiveness in this pilot phase of development and school-wide implementation. In an effort to obtain a more reliable estimate of effectiveness, therefore, we will combine these results to capture the average effect across all three SO1 schools. We do this by calculating the estimated impact at each school and taking a simple average of the estimated impacts.

Missing Data

In identifying comparison schools we created a school-level file with average test scores, and trends from 2005-06 through 2009-10 and student characteristics from 2009-10. These characteristics included the percent of the student body from different ethnic groups, the percent receiving ELL and special education services, the percent of student over age for grade, and the average attendance rate. Of the 189 New York City schools serving students in 6th to 8th grades from 2006 to 2010, we had no missing data on any of these characteristics, enabling us to find the best matching schools possible for the three SO1 schools.

Once the comparison schools were identified we created a dataset including all students that took their math tests at the eighteen comparison schools and the three SO1 schools between 2005-06 and 2010-11. See Table A5 for rates of missingness for different variables in our dataset. In the case that a test score or prior year attendance rate was missing, we mean imputed this value and included an indicator variable for missingness. In the case that a categorical variable was missing (ethnicity, ELL and special education services, overage), we dropped the observation; this accounted for 4.3 percent of our sample. Despite these techniques, our mean values changed very little (see Table A6).

Table A5
Missingness by Variable

Variable	Percent of students missing covariates (%)
Race/ethnicity	0.7
ELL services	0.5
Special education services	0.4
Free/reduced lunch status	3.5
Over age status	0.7
Gender	0.7
2010 NYS math test score	7.3
2010 Attendance rate	6.0

Source: Research Alliance analysis of student characteristics.

Table A6
Means for Complete Sample and Sample with Casewise Deletion

Characteristic	Complete Sample	Casewise Deletion
ELL status[a]		
Bilingual (%)	4	4
English as a Second Language (%)	11	12
Non-ELLs (%)	85	84
Special education status		
Small class (%)	5	5
Mainstream (%)	9	9
General education (%)	86	86
Race/ethnicity		
Black (%)	17	17
Hispanic (%)	42	43
Asian (%)	25	25
White (%)	14	15
Gender		
Female (%)	48	48
Male (%)	52	52
Grade		
6th Grade (%)	30	30
7th Grade (%)	34	34
8th Grade (%)	36	36
Over age for grade[b] **(%)**	24	24
Free/reduced lunch (%)	61	64
Past year attendance rate (%)	93	93
Average scaled score on 2010 NYS math test	664	665
Sample Size	135,474	129,766

Source: Research Alliance analysis of student characteristics.

[a] ELL designates English Language Learners.

[b] Over age for grade designates students who have repeated grades.

APPENDIX B: IMPACT ESTIMATES FOR ALL GRADES, SUBGROUPS, AND SCHOOLS

Table B1
Average Scale Scores and Estimated Differences and Standard Errors, Pooled Across Schools A, B, and C, for Grade 6 Students

Sample	SO1 Schools (Observed)	Comparison Groups (Estimated)	Difference (Estimated)	Standard Error (Estimated)
All 6th graders	671.4	671.3	0.1	1.6
Level on New York State math test in 5th Grade				
Level 1	645.3	638.0	7.4	17.1
Level 2	652.2	654.9	-2.7	4.8
Level 3	678.5	676.9	1.6	2.1
Level 4	696.7	695.5	1.2	3.7
Race/ethnicity				
Asian	680.9	673.3	7.6	11.5
Black	654.4	653.5	0.9	5.0
Hispanic	662.8	659.5	3.2	3.1
White	678.1	670.6	7.5	16.6
Gender				
Female	669.7	670.7	-1.0	2.4
Male	672.8	672.0	0.7	2.2
English as a Second Language	661.7	658.3	3.4	4.5
Mainstream Special Education[5]	656.3	660.9	-4.5	4.3

Source: Research Alliance analysis of New York State math test scores.
Note: No estimates are statistically significant. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

[5] Mainstream special education includes students with IEPs that do not require a special small classroom. These students were served by SO1 but are excluded from overall estimates of its impact.

Table B2
Average Scale Scores and Estimated Differences and Standard Errors, Pooled Across Schools A and B, for Grade 7 Students

Sample	SO1 Schools (Observed)	Comparison Groups (Estimated)	Difference (Estimated)	Standard Error (Estimated)
All 7th graders	683.0	683.1	-0.2	1.5
Level on New York State math test in 6th Grade				
Level 1	644.8	638.5	6.3	14.4
Level 2	665.0	668.1	-3.1	3.1
Level 3	685.3	687.3	-2.0	1.9
Level 4	706.1	702.1	4.1	3.3
Race/ethnicity				
Asian	692.7	691.0	1.7	2.4
Black	655.3	662.7	-7.4	4.9
Hispanic	668.2	671.9	-3.8	3.5
White	683.7	672.2	11.5	18.1
Gender				
Female	684.7	685.5	-0.8	2.2
Male	681.6	681.1	0.5	2.1
English as a Second Language	677.8	677.5	0.3	5.6
Mainstream Special Education	655.9	661.1	-5.2	3.3

Source: Research Alliance analysis of New York State math test scores.
Note: No estimates are statistically significant. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table B3
Average Scale Scores and Estimated Differences and Standard Errors, Pooled Across Schools A and B, for Grade 8 Students

Sample	SO1 Schools (Observed)	Comparison Groups (Estimated)	Difference (Estimated)		Standard Error (Estimated)
All 8th graders	685.6	689.7	-4.1	**	1.5
Level on New York State math test in 7th Grade					
Level 1	650.1	649.3	0.8		18.4
Level 2	661.1	665.7	-4.5		3.4
Level 3	685.0	689.4	-4.4	*	1.9
Level 4	713.9	715.5	-1.6		3.1
Race/ethnicity					
Asian	697.4	702.4	-5.0	*	2.4
Black	666.8	667.5	-0.7		4.8
Hispanic	667.0	674.2	-7.2	+	3.8
White	N.A.				
Gender					
Female	684.3	688.8	-4.5	*	2.3
Male	686.7	690.6	-3.9	+	2.1
English as a Second Language	676.0	684.2	-8.2		5.5
Mainstream Special Education	653.6	656.2	-2.6		3.9

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table B4
Average Scale Scores and Estimated Differences and Standard Errors, For Each SO1 School, Grade 6 Students

	School A					School B					School C				
Sample	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)	S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)
All 6th Graders	682.0	672.6	9.5	**	3.3	679.3	680.3	-1.0	2.3	652.8	660.7	-7.9	**	2.7
Level on New York State math test in 5th Grade														
Level 1	657.0	638.1	18.9		41.4	643.3	633.4	9.9	25.7	635.8	642.5	-6.7		16.2
Level 2	655.9	645.5	10.5		11.7	654.7	657.3	-2.6	6.9	645.9	661.9	-16.0	**	5.0
Level 3	681.0	667.3	13.6	**	4.4	684.7	686.3	-1.6	2.9	669.8	677.1	-7.3	*	3.6
Level 4	698.0	697.0	1.1		5.8	710.1	708.9	1.2	4.4	681.9	680.8	1.2		8.3
Race/ethnicity														
Asian	689.0	681.9	7.0	+	4.0	695.8	698.3	-2.5	4.6	658.0	639.6	18.4		34.1
Black	656.4	637.3	19.1		13.0	655.0	662.8	-7.8	6.2	651.8	660.4	-8.6	+	4.4
Hispanic	662.4	648.5	13.9	+	7.7	672.7	669.0	3.8	4.1	653.1	661.1	-8.0	*	3.4
White	688.0	663.9	24.1		33.0	678.2	679.9	-1.7	4.5	N.A.				
Gender														
Female	680.0	674.8	5.2		5.0	676.5	678.2	-1.7	3.3	652.8	659.2	-6.4	+	3.8
Male	684.0	671.3	12.7	**	4.4	681.4	682.1	-0.7	3.2	652.9	662.7	-9.8	**	3.8
ESL	678.1	667.2	10.9	+	6.4	664.6	664.3	0.4	8.9	642.4	643.5	-1.1		7.7
Mainstream	665.9	654.1	11.8	+	6.3	663.1	673.3	-10.2	6.2	639.9	655.1	-15.2		9.3

Source: Research Alliance analysis of New York State math test scores.

Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table B5
Average Scale Scores and Estimated Differences and Standard Errors, For Each SO1 School, Grade 7 Students

	School A					School B				
Sample	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)
All 7th graders	686.0	682.1	3.9	+	2.3	679.9	684.1	-4.3	*	1.9
Level on New York State math test in 6th Grade										
Level 1	649.3	638.1	11.2		24.1	640.3	638.9	1.4		15.7
Level 2	669.3	666.2	3.0		4.5	660.7	670.0	-9.3	*	4.2
Level 3	686.3	686.7	-0.5		2.9	684.3	687.9	-3.6		2.4
Level 4	710.4	700.1	10.3	*	5.0	701.9	704.0	-2.1		4.1
Race/ethnicity										
Asian	691.9	687.5	4.5	+	2.7	693.5	694.6	-1.1		3.9
Black	648.5	654.1	-5.6		8.5	662.1	671.2	-9.1	+	4.9
Hispanic	665.9	667.4	-1.5		6.0	670.5	676.5	-6.0	+	3.6
White	689.0	662.7	26.3		35.9	678.5	681.7	-3.2		3.8
Gender										
Female	690.3	686.2	4.2		3.4	679.0	684.8	-5.8	*	2.8
Male	682.5	678.5	4.0		3.1	680.7	683.6	-2.9		2.7
ESL	689.5	684.4	5.1		5.7	666.1	670.6	-4.4		9.5
Mainstream	658.3	655.8	2.5		4.5	656.1	671.0	-14.9	*	6.0

Source: Research Alliance analysis of New York State math test scores.

Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table B6
Average Scale Scores and Estimated Differences and Standard Errors, For Each SO1 School, Grade 8 Students

	School A					School B			
Sample	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)	S.E. (Est)
All 8th graders	686.2	692.8	-6.6	**	2.4	685.0	686.6	-1.6	1.9
Level on New York State math test in 7th grade									
Level 1	664.6	673.4	-8.8		33.3	635.5	625.3	10.3	15.7
Level 2	665.5	672.3	-6.8		5.5	656.8	659.1	-2.2	4.0
Level 3	687.5	693.7	-6.2	*	2.9	682.5	685.2	-2.6	2.4
Level 4	711.8	716.7	-4.9		5.0	716.0	714.3	1.7	3.7
Race/ethnicity									
Asian	692.5	698.8	-6.3	*	2.9	702.3	705.9	-3.7	3.9
Black	658.1	660.5	-2.4		8.5	675.5	674.5	1.0	4.2
Hispanic	664.0	672.6	-8.6		6.4	670.1	675.9	-5.8	4.3
White	N.A.					681.0	682.7	-1.7	3.5
Gender									
Female	686.1	691.8	-5.8	+	3.5	682.5	685.7	-3.3	2.9
Male	686.3	693.7	-7.4	*	3.4	687.1	687.5	-0.4	2.5
ESL	688.9	693.0	-4.1		6.4	663.1	675.5	-12.4	9.0
Mainstream	658.8	675.3	-16.5	***	4.8	663.6	657.8	5.8	7.0

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001.
Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

APPENDIX C: SENSITIVITY TESTS OF SAMPLE SPECIFICATION

Table C1
Average Scale Scores and Estimated Differences and Standard Errors, Pooled Across Schools A, B, and C, for Grade 6 Students that SO1 Reports Participated at Least 70% of the Time

Sample	SO1 Schools (Observed)	Comparison Groups (Estimated)	Difference (Estimated)	Standard Error (Estimated)
All 6th Graders	677.1	676.6	0.5	1.7
Level on New York State math test in 5th Grade				
Level 1	651.7	638.5	13.2	18.6
Level 2	655.9	657.0	-1.1	5.3
Level 3	678.5	677.1	1.3	2.2
Level 4	696.8	695.3	1.5	3.8
Race/ethnicity				
Asian	681.9	681.5	0.4	9.2
Black	662.9	663.9	-1.0	5.5
Hispanic	667.6	663.7	3.8	3.3
White	687.6	685.8	1.8	28.6
Gender				
Female	675.5	675.7	-0.1	2.5
Male	678.5	677.5	1.0	2.3
English as a Second Language	667.0	661.9	5.2	4.9
Mainstream Special Education[6]	665.5	665.6	-0.1	4.9

Source: Research Alliance analysis of New York State math test scores.
Note: No estimates are statistically significant. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

[6] Mainstream special education includes students with IEPs that do not require a special small classroom. These students were served by SO1 but are excluded from overall estimates of its impact.

Table C2
Average Scale Scores and Estimated Differences and Standard Errors, Pooled Across Schools A, B, and C, for Grade 7 Students that SO1 Reports Participated at Least 70% of the Time

Sample	SO1 Schools (Observed)	Comparison Groups (Estimated)	Difference (Estimated)		Standard Error (Estimated)
All 7th Graders	674.8	676.6	-1.8		1.7
Level on New York State math test in 6th Grade					
Level 1	648.7	645.7	3.0		11.1
Level 2	661.4	667.5	-6.1	*	2.8
Level 3	688.4	682.5	5.9		3.3
Level 4	707.1	704.3	2.8		3.3
Race/ethnicity					
Asian	684.1	683.1	1.1		2.4
Black	665.4	664.2	1.2		4.3
Hispanic	662.5	669.2	-6.7	*	3.0
White	N.A.				
Gender					
Female	671.0	677.5	-6.5	*	2.6
Male	677.4	675.3	2.1		2.3
English as a Second Language	669.7	671.8	-2.1		4.6
Mainstream Special Education	654.3	662.6	-8.3	+	4.2

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table C3
Average Scale Scores and Estimated Differences and Standard Errors, Pooled Across Schools A and B, for Grade 8 Students that SO1 Reports Participated at Least 70% of the Time

Sample	SO1 Schools (Observed)	Comparison Groups (Estimated)	Difference (Estimated)		Standard Error (Estimated)
All 8th Graders	690.2	693.0	-2.7	+	1.6
Level on New York State math test in 7th Grade					
Level 1	658.8	640.4	18.4		20.8
Level 2	662.6	667.8	-5.2		3.5
Level 3	686.9	690.2	-3.3	+	1.9
Level 4	714.3	715.5	-1.2		3.1
Race/ethnicity					
Asian	699.5	703.4	-3.9		2.5
Black	670.5	672.5	-2.0		5.0
Hispanic	671.5	676.5	-5.0		3.9
White	N.A.				
Gender					
Female	689.7	692.9	-3.2		2.3
Male	690.6	693.3	-2.7		2.1
English as a Second Language	687.6	688.0	-0.4		7.1
Mainstream Special Education	663.3	668.5	-5.2		4.5

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table C4
Average Scale Scores and Estimated Differences and Standard Errors,
For Each SO1 School, Grade 6 Students that SO1 Reports Participated at Least 70% of the Time

	School A					School B					School C				
Sample	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)
All 6th Graders	683.6	674.7	8.9	**	3.4	685.3	685.9	-0.6		2.4	662.5	669.3	-6.8	*	3.0
Level on New York State math test in 5th Grade															
Level 1	659.0	656.7	2.3		41.3	648.0	616.5	31.5		30.5	648.1	642.2	5.9		21.5
Level 2	659.2	648.1	11.1		13.0	657.4	658.8	-1.4		7.3	651.0	664.0	-13.0	*	5.6
Level 3	680.8	667.5	13.3	**	4.4	685.0	686.3	-1.3		3.0	669.7	677.6	-7.9	*	3.7
Level 4	698.2	696.6	1.6		5.9	710.1	708.8	1.3		4.4	682.2	680.6	1.6		8.9
Race/ethnicity															
Asian	689.6	682.6	7.0	+	4.0	698.2	701.2	-3.0		4.6	658.0	660.8	-2.8		27.0
Black	659.3	648.0	11.4		14.3	664.6	672.9	-8.3		6.8	664.8	670.8	-5.9		4.8
Hispanic	665.1	651.3	13.8	+	8.2	676.5	671.6	4.9		4.3	661.1	668.3	-7.2	+	3.8
White	688.0	681.2	6.8		56.9	687.2	688.6	-1.4		4.6	N.A.				
Gender															
Female	679.6	674.2	5.5		5.1	684.2	685.3	-1.1		3.5	662.8	667.5	-4.7		4.3
Male	687.3	675.7	11.6	*	4.6	686.1	686.7	-0.6		3.2	662.2	670.3	-8.1	*	4.1
ESL	678.1	667.2	10.8	+	6.4	666.8	664.3	2.5		9.1	656.2	654.1	2.1		9.5
Mainstream	670.1	654.8	15.3	*	6.4	668.0	672.9	-4.9		6.5	658.4	669.1	-10.7		11.5

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table C5
Average Scale Scores and Estimated Differences and Standard Errors,
For Each SO1 School, Grade 7 Students that SO1 Reports Participated at Least 70% of the Time

	School A					School B					School C				
Sample	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)
All 7th Graders	687.1	682.9	4.2	+	2.3	682.6	686.1	-3.6	+	2.0	654.7	660.6	-5.9		4.2
Level on New York State math test in 6th Grade															
Level 1	665.8	635.2	30.6		25.9	638.8	634.0	4.8		16.9	641.5	667.9	-26.4	*	12.9
Level 2	670.9	667.9	3.0		4.7	661.8	669.9	-8.1	+	4.3	651.5	664.7	-13.2	*	5.5
Level 3	686.2	687.0	-0.8		2.9	685.6	688.2	-2.5		2.5	693.3	672.2	21.1	*	9.0
Level 4	711.2	701.1	10.1	*	5.0	703.0	704.6	-1.7		4.2					
Race/ethnicity															
Asian	691.6	687.1	4.5	+	2.7	694.8	696.1	-1.3		3.9	N.A.				
Black	658.0	658.4	-0.4		9.1	664.1	671.9	-7.8		5.3	674.1	662.3	11.8		7.5
Hispanic	668.0	670.8	-2.8		6.3	672.4	677.8	-5.4		3.7	647.0	658.9	-11.9	*	5.2
White	N.A.					681.7	684.0	-2.2		4.0	N.A.				
Gender															
Female	688.6	685.2	3.4		3.4	681.0	686.7	-5.7	*	2.8	643.4	660.4	-17.0	**	6.5
Male	685.8	680.9	4.9		3.2	684.1	685.6	-1.5		2.8	662.3	659.4	2.8		5.5
ESL	691.2	685.0	6.2		5.8	668.1	671.9	-3.8		9.6	649.9	658.5	-8.6		8.0
Mainstream	660.2	658.2	2.0		4.7	656.2	672.9	-16.8	**	6.2	646.7	656.7	-10.0		10.0

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table C6
Average Scale Scores and Estimated Differences and Standard Errors,
For Each SO1 School, Grade 8 Students that SO1 Reports Participated at Least 70% of the Time

	School A					School B				
Sample	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)	SO1 (Obs)	Comp. (Est)	Diff. (Est)		S.E. (Est)
All 8th Graders	689.0	694.9	-5.9	*	2.4	691.5	691.0	0.4		2.0
Level on New York State math test in 7th Grade										
Level 1	663.8	663.0	0.8		36.8	653.8	617.8	36.0	+	19.7
Level 2	667.5	674.1	-6.6		5.5	657.7	661.5	-3.8		4.4
Level 3	689.4	694.3	-4.9	+	2.9	684.5	686.1	-1.6		2.5
Level 4	711.8	716.8	-5.0		5.0	716.7	714.1	2.6		3.8
Race/ethnicity										
Asian	694.4	699.6	-5.3	+	2.9	704.6	707.2	-2.5		3.9
Black	661.8	667.6	-5.8		9.0	679.2	677.5	1.7		4.5
Hispanic	667.8	673.0	-5.2		6.4	675.1	680.0	-4.9		4.6
White	N.A.					691.3	687.1	4.2		3.7
Gender										
Female	688.7	693.7	-4.9		3.5	690.8	692.2	-1.4		3.0
Male	689.2	696.2	-7.0	*	3.3	692.0	690.4	1.7		2.7
ESL	692.5	693.7	-1.2		6.4	682.7	682.3	0.4		12.6
Mainstream	659.7	674.9	-15.2	**	4.8	666.9	662.1	4.8		7.5

Source: Research Alliance analysis of New York State math test scores.
Note: Statistical significance of estimated differences is indicated by: + = p < .10; * = p < .05; ** = p<.01; *** = p<.001. Estimates are regression adjusted to control for differences between SO1 and comparison schools due to individual student characteristics (including race/ethnicity, gender, English Language Learner, special education, Free Lunch and holdover status, age, and prior test scores and attendance) and school-level trends in math achievement from 2006 to 2010.

Table C7
Number of Students in the Above Results Tables (A13-A18)
as Compared to the Main Results Tables (A7-A12)

		Main Results Tables	Excluded Students	Included Students
School A:	Grade 6	81	7	0
	Grade 7	149	4	0
	Grade 8	178	8	0
School B:	Grade 6	194	61	0
	Grade 7	207	39	0
	Grade 8	262	147	0
School C:	Grade 6	184	63	0
	Grade 7	0	0	32
TOTAL:		1255	329	32

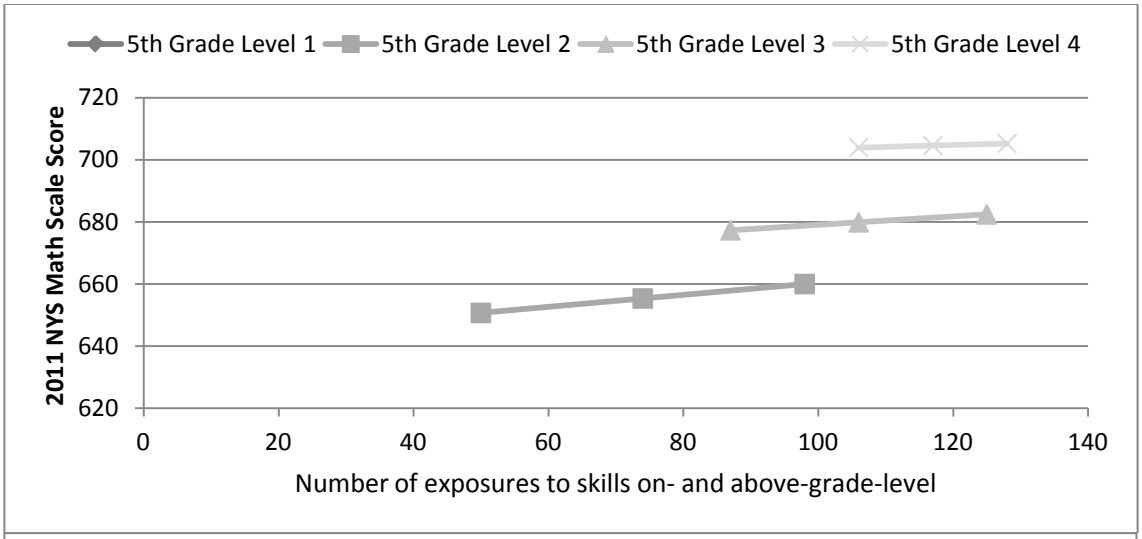
SO1 program officers asked us to exclude students for the following reasons:

Reason	% Students	Description
Plan 1	31%	Students were pulled from SO1 programming to participate in "Plan 1" for at least 30% of the scheduled SO1 instructional days. "Plan 1" was developed for students with extremely weak math skills—below fourth grade level—that SO1 could not serve.
Chronic absentees	45%	For students that did not attend SO1 at least 30% of scheduled SO1 instructional days, SO1 program officers do not consider these students have received a sufficient dosage of SO1 instruction to have a meaningful impact on student test scores.
Grade 8 SP at School B	24%	Three sections of grade 8 students at School B only received SO1 instruction as a supplement (3 periods a week) to their normal math instruction (5 periods a week). SO1 program officers do not consider these students have received a sufficient dosage of SO1 instruction to have a meaningful impact on student test scores.

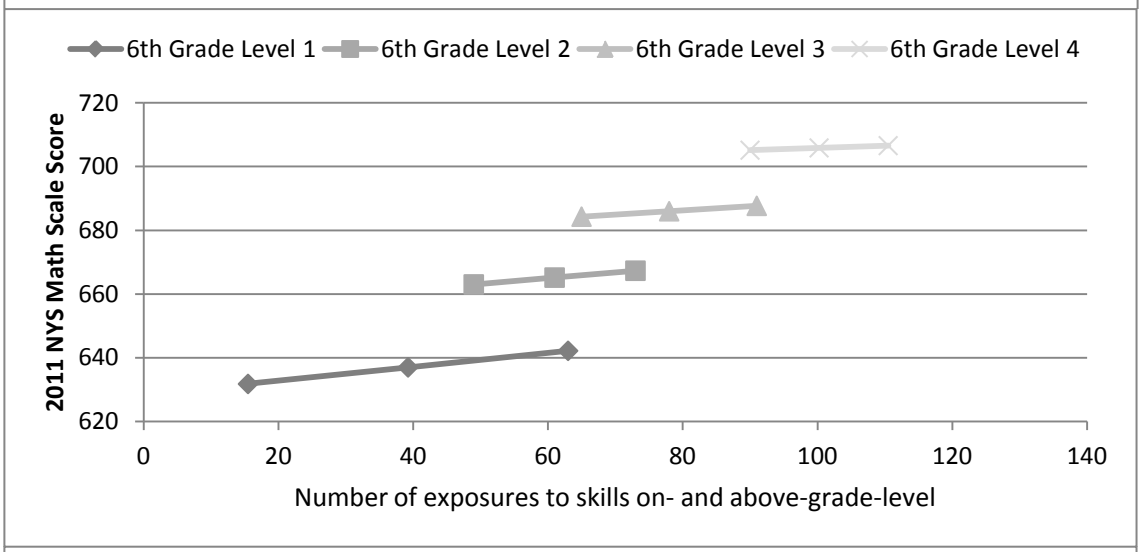
SO1 program officers asked us to include two classes of grade 7 students at School C that were given SO1 instruction, unlike most of the grade 7 students at this school. Of these 61 students, many spent at least 30% of the SO1 instructional days in "Plan 1", but 32 did not and we include these in the above results tables.

APPENDIX D: ASSOCIATION BETWEEN ON-GRADE-LEVEL EXPOSURES AND TEST SCORE GROWTH BY 2010 NYS MATH TEST PERFORMANCE LEVEL

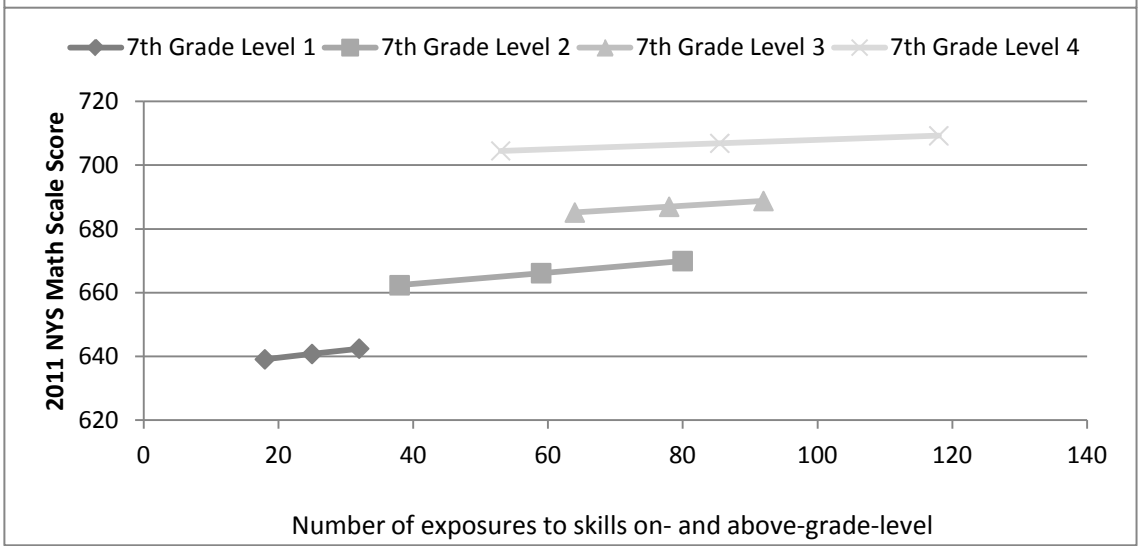
6th Grade



7th Grade



8th Grade



Source: Research Alliance analysis of SO1 internal data and New York State math test scores.

The Research Alliance for New York City Schools

285 Mercer Street, 3rd Floor | New York, New York 10003-9502
212 992 7697 | 212 995 4910 fax
research.alliance@nyu.edu | www.steinhardt.nyu.edu/research_alliance

The Research Alliance for New York City Schools conducts rigorous studies on topics that matter to the city's public schools. We strive to advance equity and excellence in education by providing non-partisan evidence about policies and practices that promote students' development and academic success.