Iowa Math Outcomes Analysis 2016/17

Grade Levels: 3, 4, 5
ST Math Program: Gen-4
Analysis Type: Z-score of scale score
Treatment-Years: 2015/16 and 2016/17
Baseline-Year: 2014/15
Subgroup: All Students

Jessica Guise
MIND Research Institute

Abstract

This analysis evaluates grades using ST Math in Iowa in 2016/17. It identifies those grades with nominal or better implementation of the ST Math program, and matches them to randomly selected, similar math-performance comparison grades. The nominal ST Math users are an aggregation of 31 grades, consisting of grades 3, 4, and 5 at 15 schools. Refer to Figures 2 and 3 for the math performance and demographic distributions. They were matched to 31 similar, randomly selected control grades at 29 schools that never used ST Math. Grade-wise growth in math proficiency was evaluated (i.e. growth in same grade, same school, from 2014/15 to 2016/17) on the percentage proficient, scale scores, and Z-scores of the scale scores (see Section 3.1). Grades 3, 4, and 5 aggregated showed an ST Math effect of 1.2 points at the Proficient or Advanced levels and Z-score of 0.04.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

This is a quasi-experimental analysis at the grade-mean level. Entire grades represent the units of analysis, and outcome measures are the 2-year changes in grade-mean ITBS Proficient or Advanced percentages. The treatment grades used the ST Math program for 2 years, beginning in the 2015/16 school year. The study hypothesis is treatment grades using ST Math will outperform similar matched control grades, using their "business as usual" conditions of instructional content and professional development. The control grades were selected to have similar demographic and math attributes (See Figures 2 and 3) to the treatment grades during the baseline year (2014/15), and did not use ST Math in 2015/16 or 2016/17. The treatment grades' selection pool was all schools using ST Math in grades 3, 4, and 5 in Iowa. The control grades' pool was all schools not using ST Math in grades 3, 4, and 5 in Iowa. This study method measures effectiveness of the ST Math program when nominally implemented.

1.2 Program Description

The ST Math program is a supplemental math program covering grade-level Iowa math standards. The ST Math content consists of visual representations of math standards, concepts, and procedures, presented to students as "Puzzles" of virtual manipulatives, with which they interact to pose solutions. Each time the student poses a solution, the computer visually animates the Puzzle, diagram, or symbols to show why the posed solution correctly solves, or why it does not solve, the math problem (puzzle). The Puzzles are arranged into sequential groups, called "Levels". To proceed to the next Level in sequence, the student needs to master his/her current Level. Mastering a Level requires solving 100% of the math problems, or Puzzles correctly. In this way, the program is self-paced. Students must correctly solve approximately 4-12 Puzzles, with only 1 failure and retry allowed, to proceed. Levels are sequenced together into Games and, again, the student must master each Game to get to the next Game in sequence. Games are sequenced into "Learning Objectives" (e.g. 'Fractions Concepts'). The ST Math curriculum of approximately 20-25 Learning Objectives can be rearranged in a year-long, grade-level syllabus to match district math pacing through the school year.

The Puzzles typically start with concrete representations of the math, without abstract symbols, math vocabulary, or even English words. Gradually, through subsequent Levels or Games, abstractions are introduced. For example, a Puzzle might start with "n" green blocks on the screen, and then at a subsequent Level may represent the quantity with the numeral for "n" (no green blocks anymore). In this way, three things are accomplished: i) language proficiency prerequisites to engage with the program are minimal, ii) non-mathematical distractions (e.g. back-stories for word problems) are minimized or eliminated – thereby reducing load on working memory, and iii) the actual math in the problem can be represented clearly, simply, and unambiguously.

Besides the self-paced progress made by students in their one-to-one environment, the program is designed to be referenced by teachers during their regular math instruction. It is supplemental to core or basal math instruction and instructional materials. As the great majority of grade-level math standards are covered in the ST Math digital curriculum, completion of 100% of the entire ST Math curriculum (i.e. completing every Game) is required to cover all grade-level math standards.

Teachers receive initial training, either face to face or through self-guided online instruction. The training covers account startup, as well as math learning and growth mindset goals, the pedagogical

approach to learning in a visual experiential game, monitoring and intervention of the student 1:1 game play, and connecting of ST Math content to classroom content and pacing.

To achieve nominal progress through the program, there is a time-on-task requirement. While student progress rates through the program vary, MIND Research Institute has found that consistent application of 90 minutes per week throughout the school year is sufficient to get most students through at least half of the ST Math Learning Objectives. Students are recommended to use the program in school for at least two 45-minute sessions per week, or 90 minutes per week, over about 35 weeks. Analyses of ST Math usage have shown that consistently following this schedule throughout the school year is usually sufficient to achieve 50% or more Progress through ST Math content. Progress is a percentage of ST Math content coverage, and is defined as Levels completed by the student, divided by the total number of Levels in the curriculum. In addition, MIND's historical analyses have shown that it is necessary to complete at least 50% of the program in order to expect significantly higher performance compared to non-users.

2 Data Collection

Since this analysis uses grades as the unit of analysis, and states publish grade-mean state standardized test scores, the data for student math outcomes is collected from each state education agency's research files (retrieved from state websites). The treatment students use ST Math student accounts served by MIND. Student ST Math usage data is aggregated to grade-level means by MIND.

2.1 Proficiency Levels Definition

The following (Table 1) is Iowa's proficiency level descriptions:

Proficiency Level	State Proficiency Level Name
L1	Not Proficient
L2	Proficient
L3	Advanced

Table 1: Proficiency Level Naming

2.2 Treatment Grades Pool and Selection

The Treatment grades pool originated with all schools and grades using ST Math in Iowa. From these schools, every grade that had used the ST Math program for the years 2015/16 and 2016/17 was identified. They comprise the Treatment grades pool for this evaluation of 3-year usage.

2.2.1 Enrollment Filter

Because the analysis uses grade-mean data, such as grade-mean scale scores or grade-mean proficiency level percentages, it is necessary that the program also be a grade-wide treatment, with the great majority of students in each grade receiving treatment. Otherwise, the grade-means reported by the state of 100% of *tested* students would not be valid measures of a smaller fraction of *treatment* students. MIND's site implementation requirement is that an entire grade, including all teachers

and all classes within that grade, use the ST Math program. We validate how closely this is the case for each individual treatment grade by comparing the number of ST Math student accounts at a grade level to the Iowa's reported enrollment at that grade level. We discard from the Treatment pool any grade with a ratio of ST Math student accounts to reported grade enrollment lower than 85%.

2.2.2 Content Coverage Filter

Furthermore, the outcomes measure is a summative year-end test, i.e. Iowa's standardized math assessment (ITBS). The math assessment thus covers all the math standards for that entire grade level. Meanwhile, the ST Math program curriculum (arranged into Learning Objectives) is also aligned to Iowa math standards. To infer that the ST Math content is having a valid effect on student outcomes on the summative assessment, we discard any grade with grade-mean of ST Math Progress for its students lower than 50% by year-end.

Progress is a percentage, and is defined as Levels completed by the student, divided by the total number of Levels in the grade-level curriculum. Note that student achievement of at least 50% progress in ST Math is accomplished primarily by teacher assignment of computer session time to students. With sufficient time on task, students make progress. The program helps them self-pace through providing real-time informative feedback for each puzzle.

2.3 Control Grades Pool and Selection

The control grades are randomly selected from a control pool of schools in Iowa. Though they are randomly selected, they are also matched to be similar to the Treatment grades' math attributes and demographics during the baseline 2014/15 year. The matched attributes include:

- scale score
- student percentages at each math proficiency level
- percentage of students receiving free or reduced lunch (using the demographic data from MDR).

To mitigate the risk of randomly picking a set of Control grades that generates an outlier for effect, a Monte Carlo approach is used to perform many random picks. The control pool's size is large enough that there are many possible "picks" of closely matched control grades.

One hundred randomly matched picks are made and sets of matched control grades are generated. For each set, the quality of the match as well as the math growth of the potential control set is evaluated. Some picked sets have high average math growth, some have low average math growth. From the set of all picks, a median pick is chosen. This avoids either an unlikely overestimate, or underestimate, of the Control grades' growth. When multiple median picks exist, the control set with the minimal math score differences in the baseline year is chosen.

3 Data Analysis

The set of all schools and grades using ST Math in Iowa is evaluated for Enrollment percentage and Progress percentage parameters. A filtered Treatment set (TRT) of all ST Math grades with $\geq 85\%$ Enrollment and $\geq 50\%$ Progress is identified. State math assessment data is tabulated. A matching set of Control grades based on baseline year state math assessment is selected.

Changes in math performance, i.e. the difference in math performance of a grade from a baseline year to the final year, are evaluated and tabulated. Statistical tests of the significance of the difference in math performance changes between Treatment grades and Control grades are performed. Finally, a grade-by-grade disaggregation is performed.

3.1 Z-scores

When states change their state assessment throughout the years, they also change the range of possible scale scores achieved on the exam. This makes it difficult to compare changes in grade mean scale scores across years with a different exam. To deal with this issue, a new Z-score is calculated. For each year being analyzed, by grade, a Z-score takes the difference of the grade mean scale score and the mean of all scale scores statewide for that year, and then divides it by the standard deviation of all scale scores statewide for that year. Here is a fictional example to illustrate the calculation of a Z-score for the 2015/16 exam:

School A, Grade 3, Mean scale score: 300
Average across all schools statewide, Grade 3: 350
Standard deviation across all schools statewide, Grade 3: 30
Z-score=((School A, Grade 3, Mean scale score)-(Average across all schools, Grade 3))/(Standard deviation across all schools, Grade 3)

$$\text{Z-score}=\frac{300-350}{30}=-1.67$$

The Z-score is calculated for every grade across all years being analyzed, using the full state data set of Iowa schools for the averages and standard deviations. The use of Z-scores is a valid statistical method to normalize any dataset and to enable analysis across otherwise uncomparable exams. In this report, we will include both mean scale scores and their accompanying Z-scores.

3.2 Percentile Ranking

These newly calculated z-scores can then be converted into a percentile ranking. Each percentile ranking shows the grade's performance relative to the others in that year and grade. For example, for a specific grade 3, a percentile ranking of 50 shows that this grade 3 performed at the average of all third grades in the state for that testing year.

3.3 Final Treatment and Control

3.3.1 ST Math Grade-Aggregated Implementation ($\geq$ 85% Enrollment Grades Only)

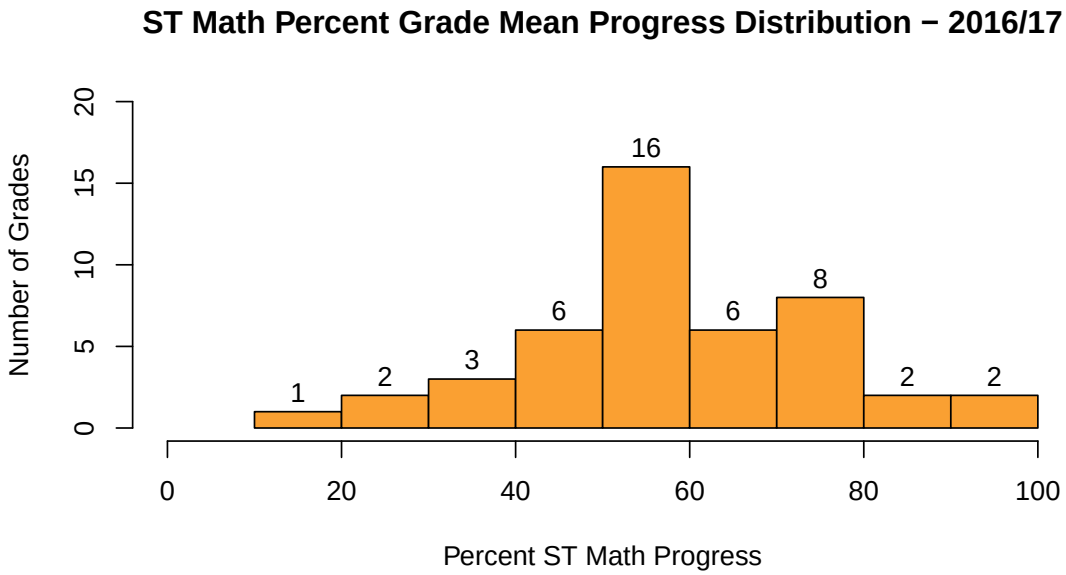

Figure 1: Histogram of ST Math Percent Progress for $\geq$ 85% Enrollment Grades 2016/17

For all ST Math grades with Enrollment $\geq$ 85%, Figure 1 shows the frequency distribution of grade-average Progress percentage through the program. Note that we will only be using grades with $\geq$ 50% Progress as the Treatment Group.

Table 2 provides descriptive statistics of the Progress distribution. Table 3 shows the number of remaining treatment grades after applying enrollment and progress filters.

	Min.	Max.	Average	S.D.
ST Math % Progress	18.8	99.0	57.7	17.3

Table 2: Descriptive Statistics of ST Math Percent Progress for >= 85 percent Enrollment Grades

Grades with >= 85% Enrollment:	39
Grades with in addition >= 50% Progress:	31

Table 3: Number of ST Math Grades with >= 85 percent Enrollment and with >= 50 percent progress

3.3.2 Filtering Treatment and Controls

Table 4 shows the total number of grades in the Treatment pool, the number of grades that exceeded the 85% Enrollment figure, and also the 50% Progress filter. Other rows in the table indicate counts of numbers of students (2016/17 from state testing count) and counts of number of schools represented. The number of matched Control (CTRL) grades, students, and schools is also shown.

	Grade 3	Grade 4	Grade 5	Total
ST Math Using Grades	16	14	16	46
ST Math Using Schools	16	14	16	23
ST Math Students	916	725	902	2543
ST Math Grades (Enroll >= 85%)	13	13	13	39
TRT Grades (Enroll >= 85% & Prog >= 50%)	11	11	9	31
TRT Schools (Enroll >= 85% & Prog >= 50%)	11	11	9	15
TRT Students (Enroll >= 85% & Prog >= 50%)	712	602	533	1847
CTRL Grades	11	11	9	31
CTRL Schools	11	11	9	29
CTRL Students	788	552	813	2153

Table 4: Treatment Pool Filtering and Controls: Counts of Grades, Schools, and Students

3.3.3 Match of Controls to Treatment

Figure 2 shows the density plot of the baseline ITBS Math scale scores (left plot) and baseline percent students at ITBS Proficient or Advanced (right plot) for treatment grades overlayed on control grades, showing the closeness of the match obtained between Treatment and Control sets of grades in the baseline year, 2014/15.

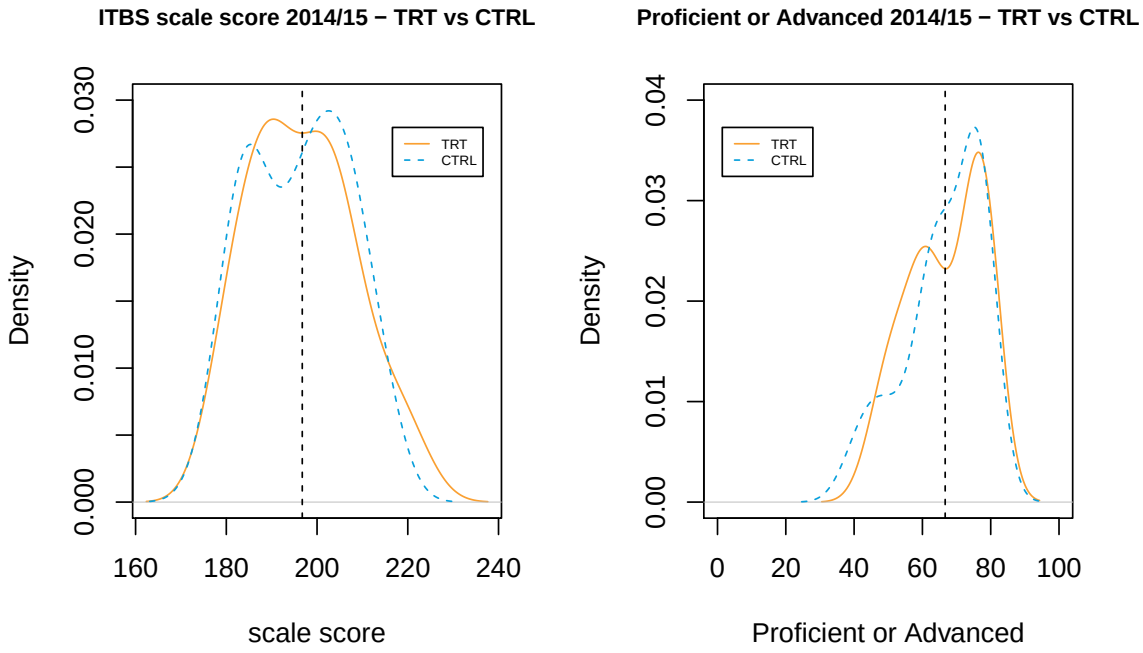

Figure 2: Baseline Year Density Plots Showing Math Scores Match between TRT and CTRL - 2014/15, with the dotted line showing the mean of the TRT set.

Similarly, figure 3 shows the density plot of the percentage of students needing free or reduced lunch for treatment grades overlayed on control grades.

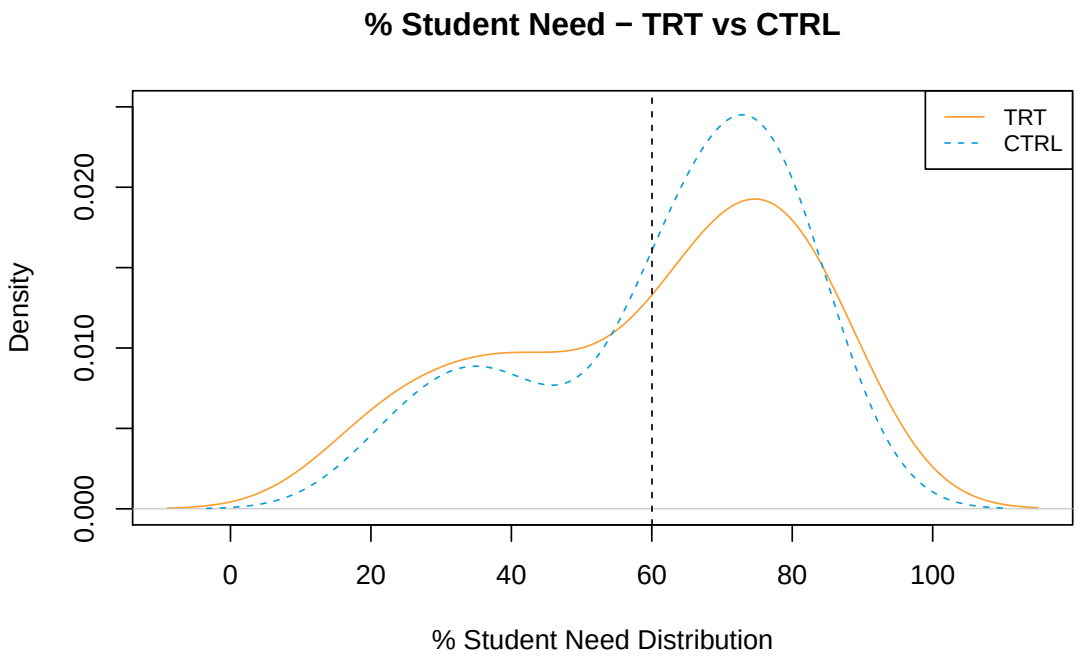

Figure 3: Baseline Year Density Plot Showing Student Need Match between TRT and CTRL, with the dotted line showing the mean of the TRT set.

Table 5 shows the difference of the means of Treatment versus Control in the baseline year, with accompanying p-values, for percent Proficient or Advanced, for mean scale score, and for percent of students receiving free or reduced lunch. The large p-values show the differences between the Treatment and Control grades are not statistically significant.

	Mean(TRT)	SD(TRT)	Mean(CTRL)	SD(CTRL)	Estimate	P-Value
Proficient or Advanced - 2014/15	66.67	10.86	66.36	11.47	0.31	0.91
Scale score - 2014/15	196.74	11.51	196.13	10.95	0.61	0.83
Percent Free or Reduced Lunch	60.03	21.33	61.81	18.69	-1.77	0.73

Table 5: Matching TRT and CTRL

3.4 Grade-Aggregated Analysis

Table 6 shows for both Treatment (TRT) and Control (CTRL) aggregation across grades of scale scores, Z-scores, and proficiency level distributions. The far right column also shows the average ST Math Progress for the TRT set.

	# Grades	# Schools	# Students	Scale score	Z-score	Percentile	L1	L2	L3	Proficient or Advanced	ST Math Per Comp.
TRT.14.15	31	15	1808	196.7	-0.70	27.74	33.33	47.95	18.71	66.67	–
TRT.15.16	31	15	1859	198.8	-0.43	36.81	30.04	47.55	22.41	69.96	59.05
TRT.16.17	31	15	1641	197.6	-0.57	32.68	32.01	47.50	20.49	67.99	63.91
TRT.Delta	–	–	–	0.8	0.14	4.94	-1.32	-0.45	1.78	1.32	–
CTRL.14.15	31	29	2320	196.1	-0.76	26.81	33.65	48.67	17.69	66.36	–
CTRL.15.16	31	29	2322	197.5	-0.57	32.68	31.30	50.72	17.98	68.70	–
CTRL.16.17	31	29	2153	196.7	-0.66	29.58	33.52	49.66	16.82	66.48	–
CTRL.Delta	–	–	–	0.6	0.10	2.77	-0.13	1.00	-0.87	0.13	–

Table 6: Yearly Math Proficiency and Counts for TRT and CTRL Grade-Aggregated Datasets

The following chart (Figure 4) shows the changes in percentage of students at each math proficiency level for the grade-aggregated Treatment and Control sets (TRT.delta and CTRL.delta).

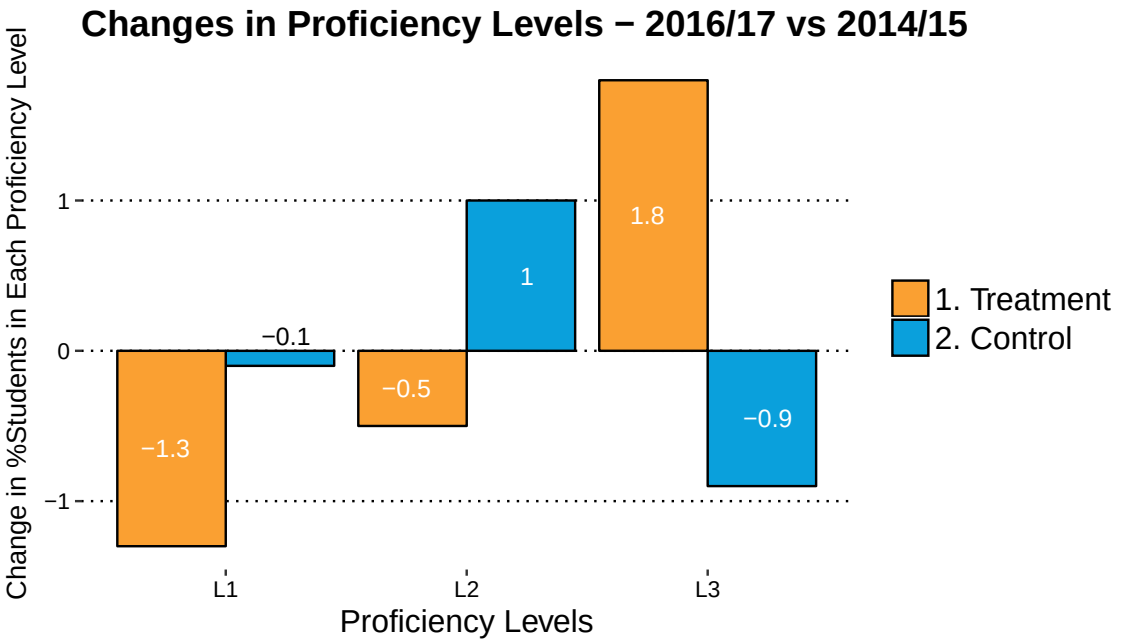

Figure 4: Change at each Proficiency Level for Grade-Aggregated TRT and CTRL Datasets between 2014/15 and 2016/17

Similarly, Figure 5 shows the changes in ITBS Math scale scores and changes in Z-scores for the grade-aggregated Treatment and Control sets.

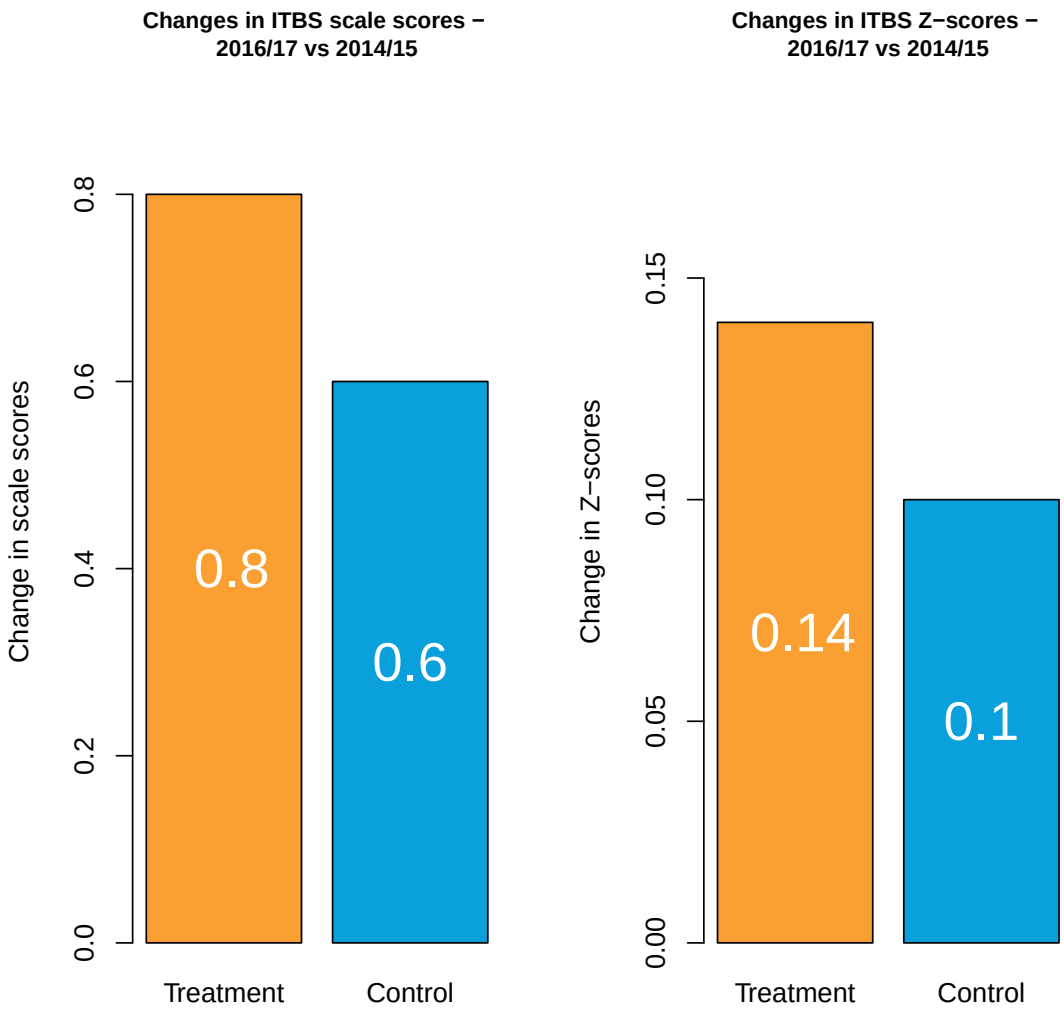

Figure 5: Changes in ITBS Math scale scores and Z-scores (See Section 3.1) for Grade-Aggregated TRT and CTRL datasets between 2014/15 and 2016/17

Further, Figure 6 shows the changes in percent of students at ITBS Proficient or Advanced for the grade-aggregated Treatment and Control sets.

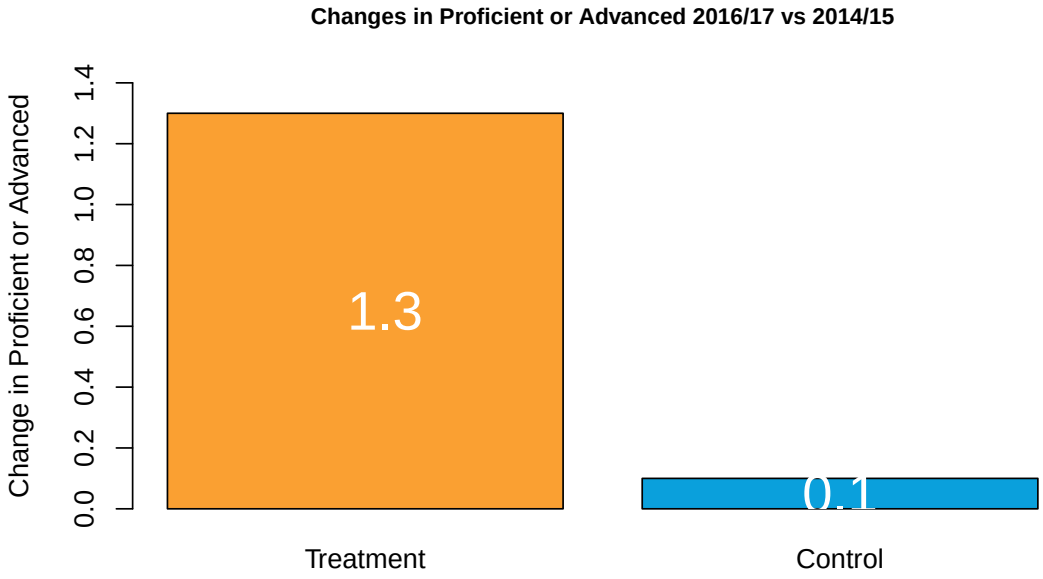

Figure 6: Changes in Proficient or Advanced for Grade-Aggregated TRT and CTRL datasets between 2014/15 and 2016/17

Finally, Table 7 shows the statistics for the *differences* in changes between TRT and CTRL (Treatment - Control) for these same ITBS math proficiency and scale score changes as in the above figures. [1]

	Estimate	P-Value	Int.Low	Int.High
Proficient or Advanced	1.20	0.66	-4.21	6.61
scale score	0.19	0.92	-3.44	3.82
Z-score	0.04	0.84	-0.33	0.41
L1	-1.19	0.66	-6.60	4.22
L2	-1.45	0.58	-6.64	3.74
L3	2.65	0.36	-3.04	8.34

Table 7: Statistics for the Differential Changes in Math Scores Growth (TRT - CTRL)

[1] * statistically significant $p<0.05$

Finally, Figure 7 shows the changes in mean percentile ranking between TRT and CTRL.

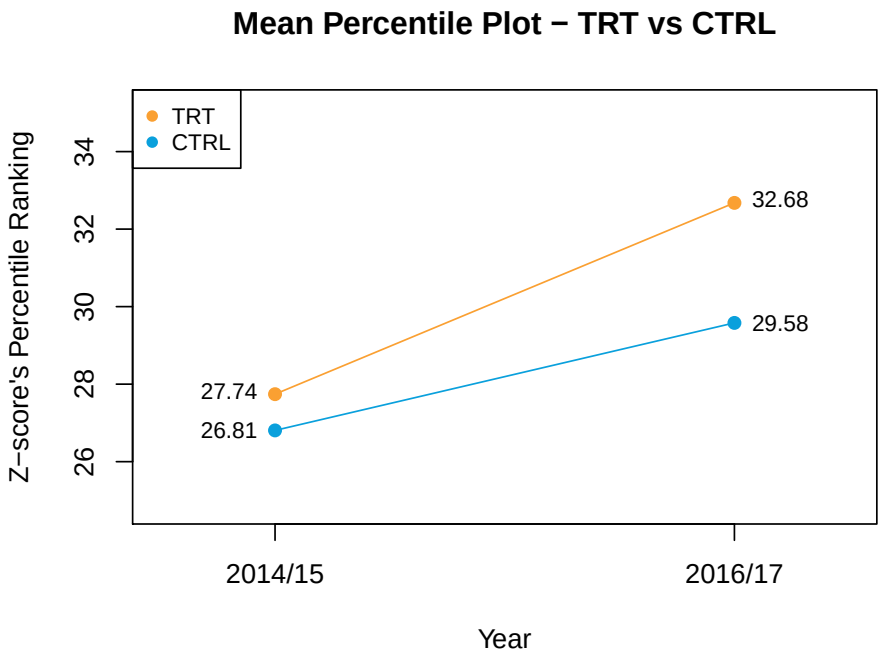

Figure 7: Changes in Percentile Ranking for TRT and CTRL Datasets between 2014/15 and 2016/17

3.5 Grade-Level Analysis

3.5.1 Grade Level Result Tables

The following tables (Table 8, 9, and 10) present a disaggregation of results by grade level. The far right column in each table also shows the average ST Math Progress for the TRT set.

	# Grades	# Schools	# Students	Scale score	Z-score	Percentile	L1	L2	L3	Proficient or Advanced	ST Math Per Prog.
TRT.14.15	11	11	695	185.2	-0.44	35.45	30.29	50.47	19.23	69.71	–
TRT.15.16	11	11	676	186.9	-0.19	44.64	29.85	46.59	23.56	70.15	60.42
TRT.16.17	11	11	610	186.4	-0.24	42.00	28.61	48.00	23.39	71.39	65.25
TRT.Delta	–	–	–	1.2	0.19	6.55	-1.68	-2.48	4.16	1.68	–
CTRL.14.15	11	11	832	184.8	-0.49	32.73	31.19	51.29	17.52	68.81	–
CTRL.15.16	11	11	844	185.7	-0.34	37.73	27.95	52.52	19.52	72.05	–
CTRL.16.17	11	11	788	183.8	-0.57	32.73	31.57	51.49	16.95	68.43	–
CTRL.Delta	–	–	–	-1.0	-0.08	0.00	0.38	0.19	-0.58	-0.38	–

Table 8: Grade 3 - Yearly Math Performance and Counts for TRT and CTRL Datasets

	# Grades	# Schools	# Students	Scale score	Z-score	Percentile	L1	L2	L3	Proficient or Advanced	ST Math Per Prog.
TRT.14.15	11	11	577	197.4	-0.82	24.00	33.18	48.13	18.69	66.82	–
TRT.15.16	11	11	614	200.6	-0.44	35.82	27.20	48.71	24.10	72.80	55.03
TRT.16.17	11	11	541	195.4	-0.98	20.73	36.39	49.34	14.27	63.61	62.81
TRT.Delta	–	–	–	-2.1	-0.16	-3.27	3.22	1.21	-4.43	-3.21	–
CTRL.14.15	11	11	643	198.6	-0.71	29.91	32.43	45.83	21.75	67.57	–
CTRL.15.16	11	11	660	199.3	-0.57	33.64	31.79	48.77	19.44	68.21	–
CTRL.16.17	11	11	552	198.6	-0.66	29.64	31.26	52.12	16.63	68.74	–
CTRL.Delta	–	–	–	0.1	0.05	-0.27	-1.17	6.29	-5.12	1.17	–

Table 9: Grade 4 - Yearly Math Performance and Counts for TRT and CTRL Datasets

	# Grades	# Schools	# Students	Scale score	Z-score	Percentile	L1	L2	L3	Proficient or Advanced	ST Math Per Prog.
TRT.14.15	9	9	536	210.0	-0.88	22.89	37.24	44.66	18.10	62.76	–
TRT.15.16	9	9	569	211.3	-0.70	28.44	33.73	47.32	18.95	66.27	62.29
TRT.16.17	9	9	490	213.9	-0.45	35.89	30.81	44.64	24.55	69.19	63.63
TRT.Delta	–	–	–	3.9	0.42	13.00	-6.43	-0.02	6.46	6.44	–
CTRL.14.15	9	9	845	207.0	-1.15	15.78	38.14	48.94	12.93	61.86	–
CTRL.15.16	9	9	818	209.7	-0.85	25.33	34.80	50.89	14.31	65.20	–
CTRL.16.17	9	9	813	210.2	-0.78	25.67	38.65	44.44	16.90	61.35	–
CTRL.Delta	–	–	–	3.2	0.37	9.89	0.52	-4.49	3.98	-0.52	–

Table 10: Grade 5 - Yearly Math Performance and Counts for TRT and CTRL Datasets

3.5.2 Grade-Level Analysis of Changes in Math Proficient or Advanced

Figure 8 shows the difference in the growth of percentages of students at math Proficient or Advanced, for the TRT and CTRL datasets, disaggregated by grade:

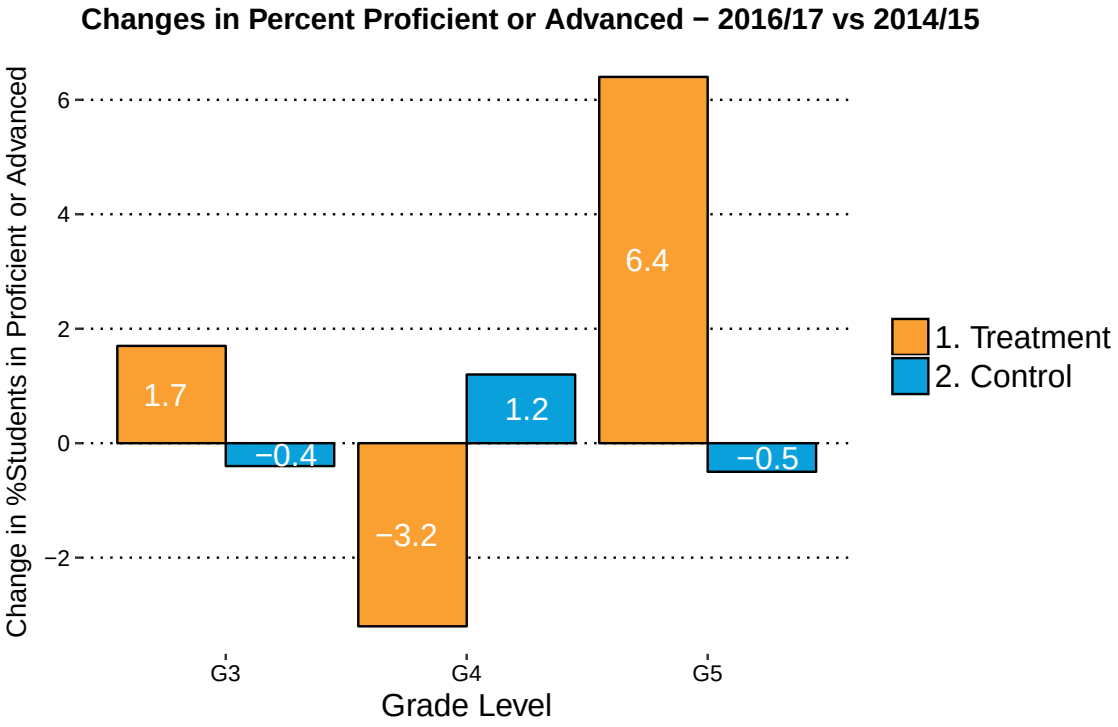

Figure 8: Changes in Percent of Students at Proficient or Advanced for TRT and CTRL Datasets between 2014/15 and 2016/17

Table 11 shows the statistics for the *differences* in changes between TRT and CTRL (Treatment - Control) for these same Proficient or Advanced math proficiency changes as shown in Figure 8.

	Estimate	P-Value	Int.Low	Int.High
Grade 3	2.07	0.57	-5.32	9.45
Grade 4	-4.38	0.44	-16.03	7.26
Grade 5	6.95	0.15	-2.83	16.74

Table 11: Statistics for the Differential Changes in Proficient or Advanced, (TRT - CTRL)

3.5.3 Grade-Level Analysis of Changes in ITBS Math scale scores

Figure 9 shows the changes in the grade-mean math scale scores of students for the TRT and CTRL datasets, disaggregated by grade:

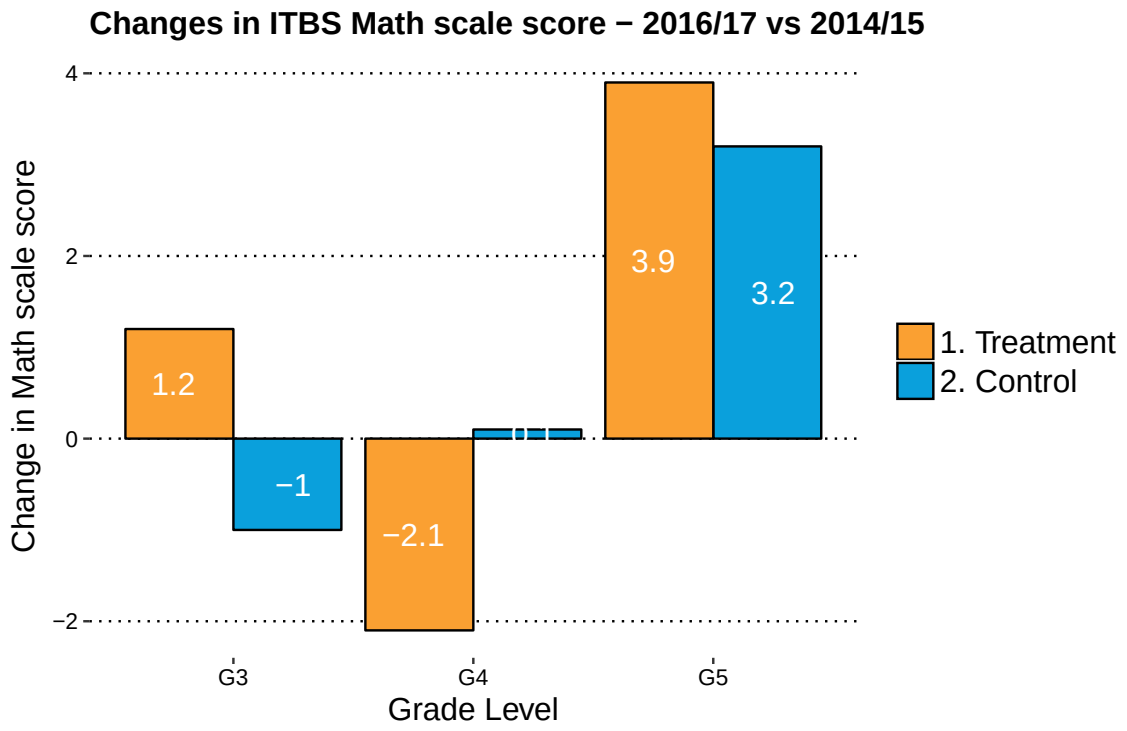

Figure 9: Changes in Grade-Mean ITBS Math scale score for TRT and CTRL Datasets between 2014/15 and 2016/17

Table 12 shows the statistics for the differences between TRT and CTRL (Treatment - Control) for these same ITBS math scale score changes as shown in Figure 9.

	Estimate	P-Value	Int.Low	Int.High
Grade 3	2.18	0.32	-2.30	6.67
Grade 4	-2.18	0.51	-8.93	4.56
Grade 5	0.67	0.87	-7.77	9.10

Table 12: Statistics for the Differential Changes in ITBS Math scale scores Growth, (TRT - CTRL)

3.5.4 Grade-Level Analysis of Changes in ITBS Z-scores of scale scores

Figure 10 shows the changes in the grade-mean Z-scores of students for the TRT and CTRL datasets, disaggregated by grade:

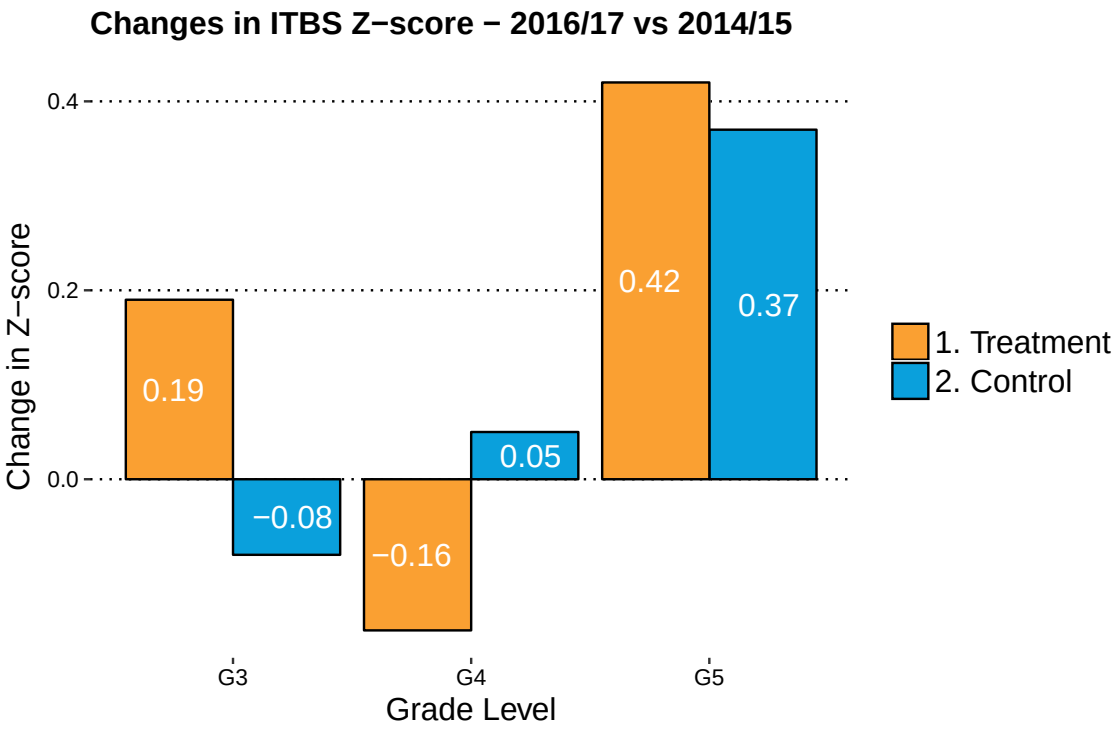

Figure 10: Changes in Grade-Mean ITBS Z-score (See Section 3.1) for TRT and CTRL Datasets between 2014/15 and 2016/17

Table 13 shows the statistics for the differences between TRT and CTRL (Treatment - Control) for these same ITBS Z-score changes as shown in Figure 10.

	Estimate	P-Value	Int.Low	Int.High
Grade 3	0.27	0.32	-0.28	0.83
Grade 4	-0.21	0.53	-0.90	0.48
Grade 5	0.06	0.88	-0.69	0.80

Table 13: Statistics for the Differential Changes in ITBS Z-scores (See Section 3.1) Growth, (TRT - CTRL)

4 Effect Size

The following table shows the effect sizes for Proficient or Advanced, ITBS scale score, and accompanying Z-score.

	Scale score Effect Size	Z-score Effect Size	Proficient or Advanced Effect Size
Grade 3	0.64	0.57	0.25
Grade 4	-0.27	-0.25	-0.36
Grade 5	0.11	0.10	0.50
All Grades	0.02	0.06	0.10

Table 14: Cohen's d Effect Size

5 Findings Summary

Iowa grades 3, 4, and 5 using ST Math for the year 2016/17 averaged 57.7% ST Math Progress. 34/46 grades (74%) averaged covering more than 50% of ST Math content. No statistically significant findings were discovered during this analysis due to the small number of treatment grades for this state.

6 Confounders

Despite best efforts in minimizing confounders to the results of this analysis, there still remain a few input variables that could be significant in affecting differences of state test score outcomes between the Treatment and Control sets. One issue is the lack of randomization of grades chosen to receive the ST Math treatment. Instead of randomized selection, Treatment grades are self-selected. Self-selection can be an indication of districts or schools with a focus on math, an appetite for change, and with a spotlight on math training. Furthermore, not all grades using the ST Math program are chosen for analysis. Each grade must pass two specific filters to be considered for the Treatment set: the first being an enrollment filter of at least 85% of students in each grade using the program, and the second being a progress filter of at least 50% of the program completed on average by students in that grade. These filters might indicate relatively high-functioning schools with a team of relatively effective teachers in that grade, thus resulting in better instruction overall. A mitigation of this possible confounder is our selection of treatment groups on the grade level, rather than the teacher level, so there is no cherry picking of teachers: the full range of teachers in each grade is included. Moreover, the specific teachers may often be the same in the baseline year as in the current year, so the Treatment growth is not due to teacher differences. Finally, a possible confounder lies in the "business as usual" conditions at the matched control grades chosen for each analysis. It's unknown whether these control grades used other programs that could affect the comparison of the two sets of grades. The Monte Carlo Method is used to mitigate the possibility of control picks being favorable or unfavorable (see Section 2.3).

7 Reference Tables Grouped By School Year

The following tables show grade-level details, grouped by school year and for treatment (Table 15) and controls (Table 16) separately.

	# Grades	# Schools	# Students	Scale score	Z-score	Percentile	L1	L2	L3	Proficient or Advanced	ST Math Per Comp.
Grade 3 (14.15)	11	11	695	185.2	-0.44	35.45	30.29	50.47	19.23	69.71	–
Grade 4 (14.15)	11	11	577	197.4	-0.82	24.00	33.18	48.13	18.69	66.82	–
Grade 5 (14.15)	9	9	536	210.0	-0.88	22.89	37.24	44.66	18.10	62.76	–
All Grades (14.15)	31	15	1808	196.7	-0.70	27.74	33.33	47.95	18.71	66.67	–
Grade 3 (15.16)	11	11	676	186.9	-0.19	44.64	29.85	46.59	23.56	70.15	60.42
Grade 4 (15.16)	11	11	614	200.6	-0.44	35.82	27.20	48.71	24.10	72.80	55.03
Grade 5 (15.16)	9	9	569	211.3	-0.70	28.44	33.73	47.32	18.95	66.27	62.29
All Grades (15.16)	31	15	1859	198.8	-0.43	36.81	30.04	47.55	22.41	69.96	59.05
Grade 3 (16.17)	11	11	610	186.4	-0.24	42.00	28.61	48.00	23.39	71.39	65.25
Grade 4 (16.17)	11	11	541	195.4	-0.98	20.73	36.39	49.34	14.27	63.61	62.81
Grade 5 (16.17)	9	9	490	213.9	-0.45	35.89	30.81	44.64	24.55	69.19	63.63
All Grades (16.17)	31	15	1641	197.6	-0.57	32.68	32.01	47.50	20.49	67.99	63.91

Table 15: TRT Grades Detail Sorted by Year

	# Grades	# Schools	# Students	Scale score	Z-score	Percentile	L1	L2	L3	Proficient or Advanced	ST Math Per Comp.
Grade 3 (14.15)	11	11	832	184.8	-0.49	32.73	31.19	51.29	17.52	68.81	–
Grade 4 (14.15)	11	11	643	198.6	-0.71	29.91	32.43	45.83	21.75	67.57	–
Grade 5 (14.15)	9	9	845	207.0	-1.15	15.78	38.14	48.94	12.93	61.86	–
All Grades (14.15)	31	29	2320	196.1	-0.76	26.81	33.65	48.67	17.69	66.36	–
Grade 3 (15.16)	11	11	844	185.7	-0.34	37.73	27.95	52.52	19.52	72.05	–
Grade 4 (15.16)	11	11	660	199.3	-0.57	33.64	31.79	48.77	19.44	68.21	–
Grade 5 (15.16)	9	9	818	209.7	-0.85	25.33	34.80	50.89	14.31	65.20	–
All Grades (15.16)	31	29	2322	197.5	-0.57	32.68	31.30	50.72	17.98	68.70	–
Grade 3 (16.17)	11	11	788	183.8	-0.57	32.73	31.57	51.49	16.95	68.43	–
Grade 4 (16.17)	11	11	552	198.6	-0.66	29.64	31.26	52.12	16.63	68.74	–
Grade 5 (16.17)	9	9	813	210.2	-0.78	25.67	38.65	44.44	16.90	61.35	–
All Grades (16.17)	31	29	2153	196.7	-0.66	29.58	33.52	49.66	16.82	66.48	–

Table 16: CTRL Grades Detail Sorted by Year

8 Lists of Schools

8.1 Treatment Schools

The following table lists the treatment schools and grades (after 85% enrollment and 50% progress filtering) used in the analysis.

PID	IID	District	School Name	GRADE
243226	JOH41K	Cedar Rapids Community School District	Johnson Elementary School	3, 4, 5
234433	CEN40K	Central Community School District	Central Elementary School	3, 4
236297	DAN426	Danville Community School District	Danville Elementary School	3
250542	FIL42O	Davenport Community School District	Fillmore Elementary School	3, 4, 5
250619	HAY42O	Davenport Community School District	Hayes Elementary School	3, 4, 5
250683	MAD42O	Davenport Community School District	Madison Elementary School	3, 4, 5
10902892	GEO40G	Dubuque Community School District	Carver Elementary School	5
244311	EAR3V3	Earlham Community School District	Earlham Elementary School	3, 4
235932	EDG40J	Edgewood-Colesburg Community School District	Edgewood-Colesburg Elementary School	4, 5
240391	CRE40T	Howard-Winneshiek Community School District	Crestwood Elementary School	5
245975	FRA42K	Muscatine Community School District	Franklin Elementary School	3, 4
234079	NEW3WQ	New Hampton Community School District	New Hampton Elementary School	3
237978	WIN3WQ	Oelwein Community School District	Wings Park Elementary School	3, 4, 5
238570	SID3ZH	Sidney Community School District	Sidney Elementary School	4, 5
246137	WES42L	West Liberty Community School District	West Liberty Elementary School	3, 4

Table 17: Treatment Schools (TRT Dataset)

8.2 Control Schools

The following table lists the control schools and grades (matched control grades to treatment grades) used in the analysis.

PID	District	School Name	GRADE
236118	Burlington Community School District	Black Hawk Elementary School	4
236182	Burlington Community School District	James Wilson Grimes School	3
243056	Cedar Rapids Community School District	Cleveland Elementary School	4
243393	Cedar Rapids Community School District	Van Buren Elementary School	4
11136606	Clear Creek Amana Community School District	North Bend Elementary	5
243537	College Community School District	Prairie View Elementary School	3
243903	Columbus Community School District	Roundy Elementary School	5
249323	Council Bluffs Community School District	Bloomer Elementary School	4
249347	Council Bluffs Community School District	Carter Lake Elementary School	3
249373	Council Bluffs Community School District	Edison Elementary School	4, 3
247777	Des Moines Independent Community School District	Brubaker Elementary School	5
248173	Des Moines Independent Community School District	South Union Elementary School	5
248331	Des Moines Independent Community School District	Stowe Elementary School	4
255994	Eagle Grove Community School District	Robert Blue School	5
244232	George-Little Rock Community School District	Little Rock Elementary School	5
241826	Iowa City Community School District	Kirkwood Elementary School	4
245080	Marshalltown Community School District	Fisher Elementary School	3
10804927	Marshalltown Community School District	Lenihan Intermediate School	5
233532	Mason City Community School District	Harding Elementary School	4
245963	Muscatine Community School District	Colorado Elementary School	3
4029917	North Union Community School District	North Union Elementary	5
254926	Sioux City Community School District	Leeds Elementary School	3
231039	Waterloo Community School District	Fred Becker Elementary School	5, 3
231156	Waterloo Community School District	Lou Henry Elementary School	3
231194	Waterloo Community School District	Lowell Elementary School	3
248824	West Des Moines Community School District	Clive Learning Academy	4
236833	Western Dubuque Community School District	Farley Elementary School	4
10010413	Western Dubuque Community School District	Peosta Elementary School	4
244402	Winterset Community School District	Winterset Elementary School	3

Table 18: Matched Control Schools (CTRL Dataset)