# Pickerington, Ohio Math Outcomes Analysis 2016/17

**Grade Levels: 3, 4, 5**
**ST Math Program: Gen-4**
**Analysis Type: Z-score of Proficient or Above**
**Treatment-Years: 2014/15, 2015/16, and 2016/17**
**Baseline-Year: 2013/14**
**Subgroup: All Students**

**Jessica Guise**

**Abstract**

This analysis evaluates grades using ST Math in Pickerington, Ohio 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 17 grades, consisting of grades 3, 4, and 5 at 10 schools. Refer to Figures 2 and 3 for the math performance and demographic distributions. They were matched to 17 similar, randomly selected control grades at 17 schools that never used ST Math. Grade-wise growth in math proficiency was evaluated (i.e. growth in same grade, same school, from 2013/14 to 2016/17) on the percentage proficient and the accompanying z-scores of the percentage proficient (see Section 3.1). Grades 3, 4, and 5 aggregated showed an ST Math effect of 4.42 points at the Proficient or Above levels, 8.48 points at the Advanced Level, and Z-score of 0.17.

# 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 3-year changes in grade-mean OAA Proficient or Above percentages. The treatment grades used the ST Math program for 3 years, beginning in the 2016/17 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 (2013/14), and did not use ST Math in 2016/17, and 2016/17. The treatment grades' selection pool was all schools using ST Math in grades 3, 4, and 5 in Pickerington, Ohio. The control grades' pool was all schools not using ST Math in grades 3, 4, and 5 in Ohio. 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 Ohio 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 Ohio's proficiency level descriptions:

| Proficiency Level | State Proficiency Level Name |
|:---:|:---|
| L1 | Limited |
| L2 | Basic |
| L3 | Proficient |
| L4 | Accelerated |
| L5 | Advanced |
| L6 | Advanced Plus |

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 Pickerington, Ohio. From these schools, every grade that had used the ST Math program for the years 2014/15, 2015/16, and 2016/17 was identified. They comprise the Treatment grades pool for this evaluation of 4-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 Ohio'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. Ohio's standardized math assessment (OAA). 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 Ohio 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 Ohio. Though they are randomly selected, they are also matched to be similar to the Treatment grades' math attributes and demographics during the baseline 2013/14 year. The matched attributes include:

- 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 Ohio 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

In order to analyze across all states with different math assessments, a new z-score of that test's math proficiency is calculated. For each year being analyzed, by grade, a z-score takes the difference of the grade mean percent proficient and the mean of all percent proficient statewide for that year, and then divides it by the standard deviation of all percent proficient 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, Percent Proficient: 70
Average across all schools statewide, Grade 3: 50
Standard deviation across all schools statewide, Grade 3: 20
Z-score=((School A, Grade 3, Percent Proficient)-(Average across all schools, Grade 3))/(Standard deviation across all schools, Grade 3)

$$\text{Z-score} = \frac{70 - 50}{20} = 1$$

The z-score is calculated for every grade across all years being analyzed, using the full state data set of 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 only analyze 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)

**ST Math Percent Grade Mean Progress Distribution – 2016/17**
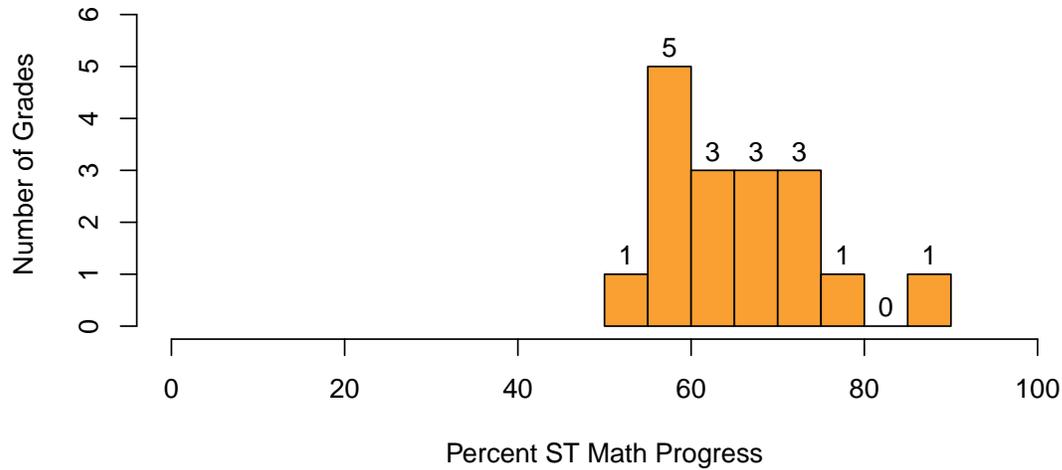


Figure 1: Histogram of ST Math Percent Progress for $\geq 85\%$ Enrollment Grades 2016/17

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 | 53.7 | 87.0 | 65.7 | 8.9 |

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

| | |
|---|---|
| Grades with $>= 85\%$ Enrollment: | 17 |
| Grades with in addition $>= 50\%$ Progress: | 17 |

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 | 7 | 7 | 3 | 17 |
| ST Math Using Schools | 7 | 7 | 3 | 10 |
| ST Math Students | 728 | 776 | 741 | 2245 |
| ST Math Grades (Enroll >= 85%) | 7 | 7 | 3 | 17 |
| TRT Grades (Enroll >= 85% & Prog >= 50%) | 7 | 7 | 3 | 17 |
| TRT Schools (Enroll >= 85% & Prog >= 50%) | 7 | 7 | 3 | 10 |
| TRT Students (Enroll >= 85% & Prog >= 50%) | 798 | 749 | 694 | 2241 |
| CTRL Grades | 7 | 7 | 3 | 17 |
| CTRL Schools | 7 | 7 | 3 | 17 |
| CTRL Students | 530 | 1064 | 620 | 2214 |

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 percent students at OAA Proficient or Above 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, 2013/14. It is important to keep in mind that we only have a small number of treatment and control grades (17) and that the Control set was arrived at through a Monte Carlo process (see Section 2.3) rather than a closest math performance match.

**Proficient or Above 2013/14 – TRT vs CTRL**
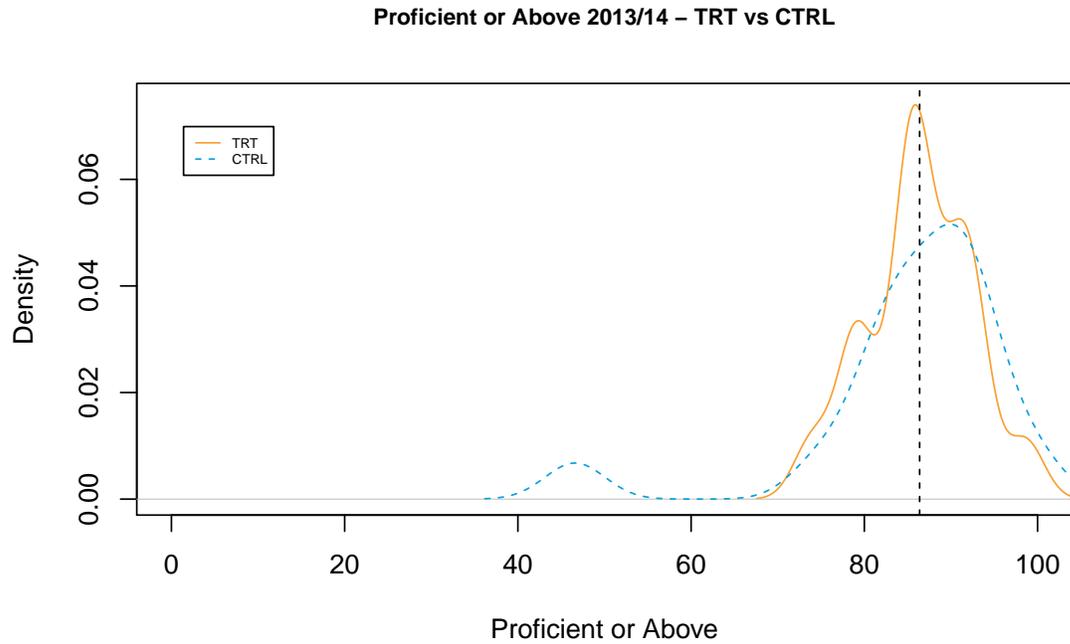


Figure 2: Baseline Year Density Plots Showing Math Scores Match between TRT and CTRL - 2013/14, with the dotted line showing the mean of the TRT set.

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

**% Student Need – TRT vs CTRL**
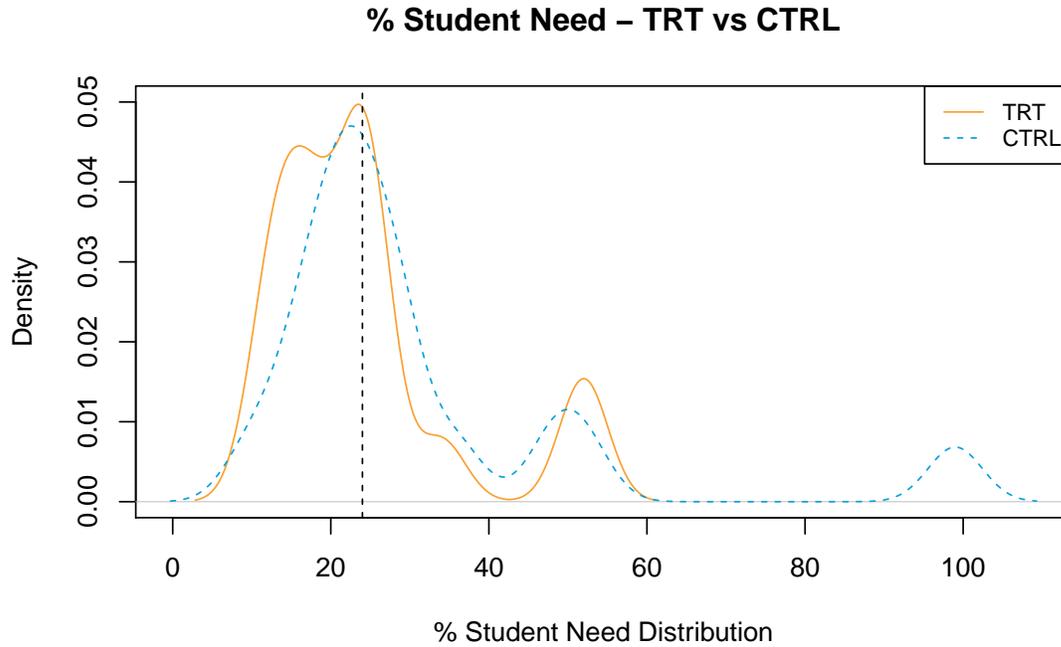


Figure 3: Baseline Year Density Plot Showing Student Need Match between TRT and CTRL, with the dotted line showing the mean of the TRT set.

Table 5 shows the difference of the means of Treatment versus Control in the baseline year, with accompanying p-values, for percent Proficient or Above 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 Above - 2013/14 | 86.38 | 6.18 | 85.51 | 11.79 | 0.87 | 0.79 |
| Percent Free or Reduced Lunch | 24.00 | 12.02 | 30.35 | 20.76 | -6.35 | 0.28 |

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 proficiency level, z-score, and percentile distributions. The far right column also shows the average ST Math Progress for the TRT set.

| | # Grades | # Schools | # Students | L1 | L2 | L3 | L4 | L5 | L6 | Proficient or Above | Z-score | Percentile | ST Math Per Comp. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TRT.13.14 | 17 | 10 | 2325 | 4.18 | 9.45 | 26.52 | 24.68 | 35.11 | 0.08 | 86.38 | 0.61 | 72.24 | – |
| TRT.14.15 | 17 | 10 | 2131 | 2.31 | 11.71 | 25.24 | 50.71 | 10.01 | 0.03 | 85.98 | 1.00 | 83.71 | 53.06 |
| TRT.15.16 | 17 | 10 | 2114 | 8.18 | 6.46 | 17.38 | 26.30 | 41.55 | 0.12 | 85.35 | 0.90 | 81.29 | 63.77 |
| TRT.16.17 | 17 | 10 | 2245 | 8.03 | 5.28 | 17.69 | 27.73 | 41.14 | 0.13 | 86.68 | 0.86 | 79.94 | 65.7 |
| TRT.Delta | – | – | – | 3.85 | -4.17 | -8.84 | 3.05 | 6.03 | 0.05 | 0.30 | 0.25 | 7.71 | – |
| CTRL.13.14 | 17 | 17 | 2243 | 5.52 | 8.94 | 28.18 | 23.26 | 33.76 | 0.31 | 85.51 | 0.57 | 70.88 | – |
| CTRL.14.15 | 17 | 17 | 2117 | 9.00 | 16.62 | 28.30 | 40.12 | 5.85 | 0.12 | 74.39 | 0.50 | 68.29 | – |
| CTRL.15.16 | 17 | 17 | 2189 | 13.13 | 9.51 | 19.00 | 27.25 | 31.04 | 0.07 | 77.36 | 0.58 | 71.00 | – |
| CTRL.16.17 | 17 | 17 | 2214 | 10.86 | 7.72 | 20.95 | 28.49 | 31.31 | 0.64 | 81.39 | 0.64 | 73.06 | – |
| CTRL.Delta | – | – | – | 5.34 | -1.22 | -7.23 | 5.23 | -2.45 | 0.33 | -4.12 | 0.07 | 2.18 | – |

Table 6: All Grades Together Growth

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).
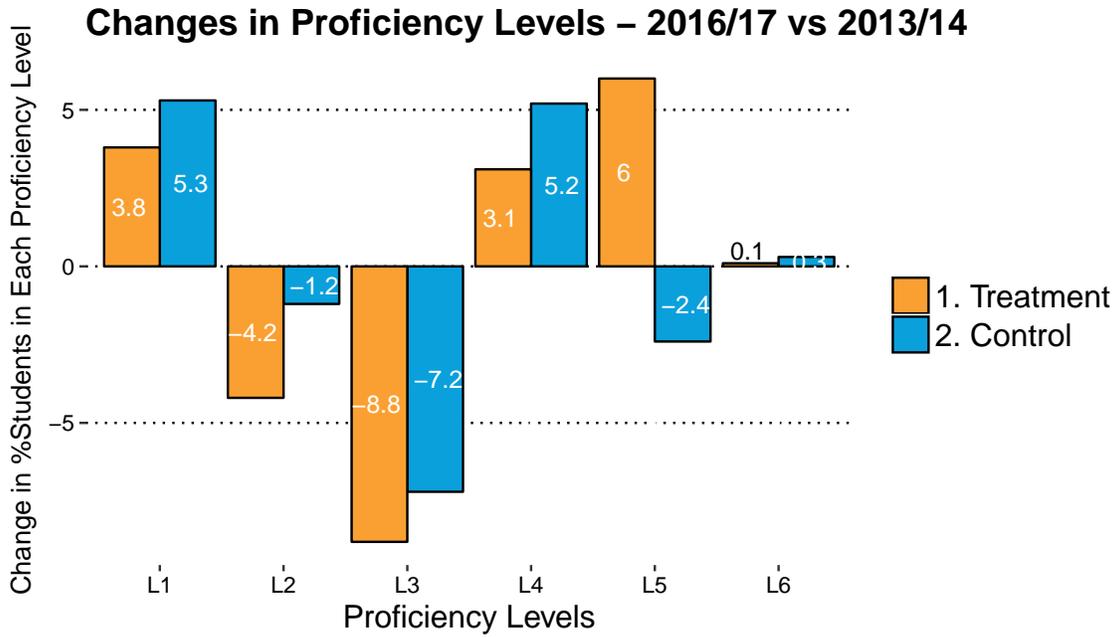


Figure 4: Change at each Proficiency Level for Grade-Aggregated TRT and CTRL Datasets between 2013/14 and 2016/17

Similarly, Figure 5 shows the changes in percent of students at OAA Proficient or Above and changes in accompanying z-scores for the grade-aggregated Treatment and Control sets.
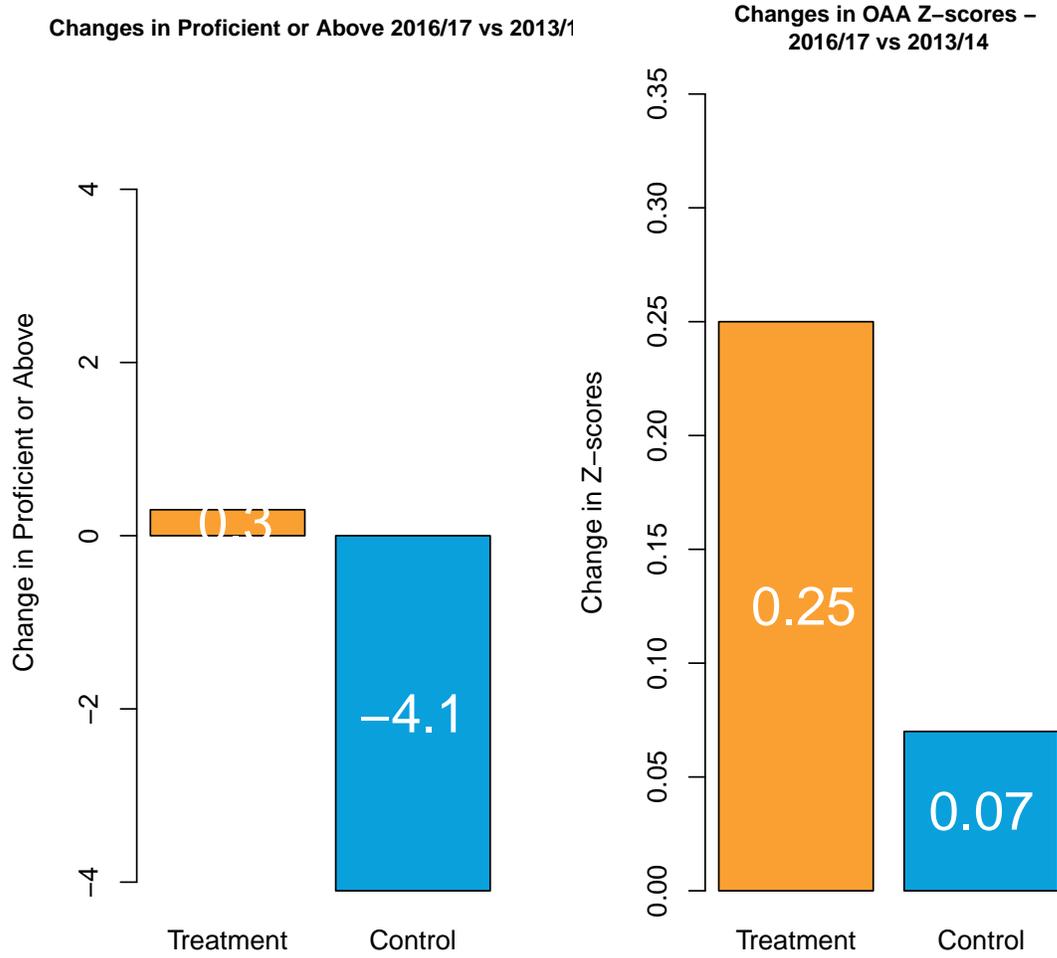


Figure 5: Changes in Proficient or Above for Grade-Aggregated TRT and CTRL datasets between 2013/14 and 2016/17

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

|  | Estimate | P-Value | Int.Low | Int.High |
|---|---|---|---|---|
| Proficient or Above | 4.42 | 0.02* | 0.76 | 8.08 |
| Z-score | 0.17 | 0.03* | 0.02 | 0.33 |
| L1 | -1.49 | 0.44 | -5.42 | 2.43 |
| L2 | -2.95 | 0.11 | -6.57 | 0.68 |
| L3 | -1.61 | 0.6 | -7.74 | 4.53 |
| L4 | -2.18 | 0.25 | -5.93 | 1.58 |
| L5 | 8.48 | 0.03* | 0.79 | 16.16 |
| L6 | -0.28 | 0.34 | -0.86 | 0.31 |

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

Finally, Figure 6 shows the changes in mean percentile ranking between TRT and CTRL.
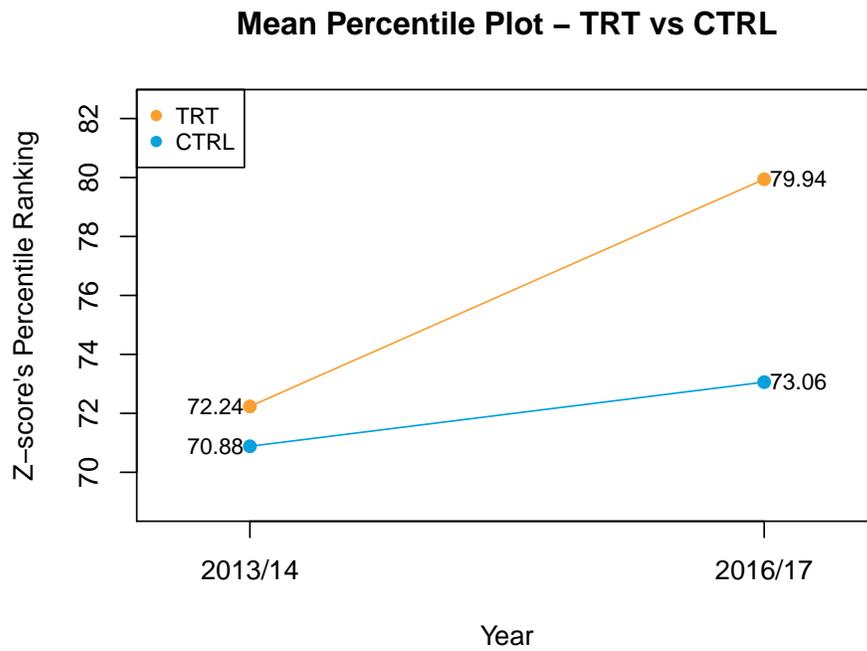


Figure 6: Changes in Percentile Ranking for TRT and CTRL Datasets between 2013/14 and 2016/17

---

[1]* statistically significant p<0.05

## 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 | L1 | L2 | L3 | L4 | L5 | L6 | Proficient or Above | Z-score | Percentile | ST Math Per Prog. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TRT.13.14 | 7 | 7 | 757 | 2.00 | 7.74 | 28.33 | 24.67 | 37.30 | 0.00 | 90.30 | 0.69 | 74.86 | – |
| TRT.14.15 | 7 | 7 | 680 | 2.84 | 10.73 | 23.13 | 49.33 | 13.97 | 0.00 | 86.43 | 0.98 | 83.00 | 56.45 |
| TRT.15.16 | 7 | 7 | 713 | 8.11 | 6.66 | 17.41 | 23.56 | 43.99 | 0.24 | 85.20 | 0.92 | 82.00 | 65.28 |
| TRT.16.17 | 7 | 7 | 728 | 7.49 | 5.89 | 17.56 | 27.64 | 41.47 | 0.00 | 86.67 | 0.84 | 79.29 | 64.06 |
| TRT.Delta | – | – | – | 5.49 | -1.86 | -10.77 | 2.97 | 4.17 | 0.00 | -3.63 | 0.15 | 4.43 | – |
| CTRL.13.14 | 7 | 7 | 524 | 2.53 | 5.16 | 28.93 | 24.04 | 38.90 | 0.43 | 92.30 | 0.81 | 78.14 | – |
| CTRL.14.15 | 7 | 7 | 555 | 4.04 | 16.31 | 26.81 | 42.71 | 9.81 | 0.30 | 79.64 | 0.67 | 73.00 | – |
| CTRL.15.16 | 7 | 7 | 537 | 11.91 | 10.66 | 18.49 | 22.20 | 36.60 | 0.17 | 77.46 | 0.59 | 70.71 | – |
| CTRL.16.17 | 7 | 7 | 530 | 5.90 | 7.93 | 17.59 | 30.06 | 37.43 | 1.07 | 86.14 | 0.82 | 78.71 | – |
| CTRL.Delta | – | – | – | 3.37 | 2.77 | -11.34 | 6.01 | -1.47 | 0.64 | -6.16 | 0.01 | 0.57 | – |

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

| | # Grades | # Schools | # Students | L1 | L2 | L3 | L4 | L5 | L6 | Proficient or Above | Z-score | Percentile | ST Math Per Prog. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TRT.13.14 | 7 | 7 | 743 | 3.97 | 10.51 | 25.17 | 28.63 | 31.71 | 0.00 | 85.51 | 0.48 | 68.00 | – |
| TRT.14.15 | 7 | 7 | 725 | 2.07 | 11.19 | 25.07 | 54.83 | 6.81 | 0.00 | 86.71 | 1.08 | 85.71 | 50.75 |
| TRT.15.16 | 7 | 7 | 664 | 8.03 | 5.57 | 13.60 | 30.41 | 42.41 | 0.00 | 86.43 | 0.83 | 79.29 | 61.61 |
| TRT.16.17 | 7 | 7 | 776 | 6.64 | 3.44 | 14.37 | 29.63 | 45.53 | 0.31 | 89.84 | 0.86 | 80.29 | 67.19 |
| TRT.Delta | – | – | – | 2.67 | -7.07 | -10.80 | 1.00 | 13.81 | 0.31 | 4.33 | 0.38 | 12.29 | – |
| CTRL.13.14 | 7 | 7 | 1095 | 6.91 | 12.00 | 29.39 | 24.97 | 26.73 | 0.00 | 81.09 | 0.25 | 61.14 | – |
| CTRL.14.15 | 7 | 7 | 1052 | 16.11 | 16.16 | 28.07 | 37.67 | 2.00 | 0.00 | 67.74 | 0.25 | 61.71 | – |
| CTRL.15.16 | 7 | 7 | 1066 | 15.64 | 7.59 | 16.27 | 33.21 | 27.29 | 0.00 | 76.77 | 0.45 | 67.86 | – |
| CTRL.16.17 | 7 | 7 | 1064 | 14.44 | 7.07 | 20.66 | 28.17 | 29.17 | 0.47 | 78.47 | 0.40 | 65.14 | – |
| CTRL.Delta | – | – | – | 7.53 | -4.93 | -8.73 | 3.20 | 2.44 | 0.47 | -2.61 | 0.15 | 4.00 | – |

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

| | # Grades | # Schools | # Students | L1 | L2 | L3 | L4 | L5 | L6 | Proficient or Above | Z-score | Percentile | ST Math Per Prog. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TRT.13.14 | 3 | 3 | 825 | 9.77 | 10.93 | 25.47 | 15.47 | 37.90 | 0.43 | 79.27 | 0.71 | 76.00 | – |
| TRT.14.15 | 3 | 3 | 726 | 1.60 | 15.20 | 30.53 | 44.30 | 8.20 | 0.17 | 83.20 | 0.89 | 80.67 | 50.54 |
| TRT.15.16 | 3 | 3 | 737 | 8.70 | 8.07 | 26.13 | 23.10 | 33.87 | 0.10 | 83.20 | 1.02 | 84.33 | 65.28 |
| TRT.16.17 | 3 | 3 | 741 | 12.53 | 8.13 | 25.73 | 23.50 | 30.10 | 0.00 | 79.33 | 0.88 | 80.67 | 66.02 |
| TRT.Delta | – | – | – | 2.77 | -2.80 | 0.27 | 8.03 | -7.80 | -0.43 | 0.07 | 0.16 | 4.67 | – |
| CTRL.13.14 | 3 | 3 | 624 | 9.27 | 10.63 | 23.63 | 17.47 | 38.17 | 0.73 | 80.00 | 0.75 | 76.67 | – |
| CTRL.14.15 | 3 | 3 | 510 | 3.97 | 18.43 | 32.30 | 39.77 | 5.57 | 0.00 | 77.63 | 0.66 | 72.67 | – |
| CTRL.15.16 | 3 | 3 | 586 | 10.10 | 11.33 | 26.57 | 25.13 | 26.83 | 0.00 | 78.53 | 0.84 | 79.00 | – |
| CTRL.16.17 | 3 | 3 | 620 | 14.10 | 8.73 | 29.50 | 25.60 | 22.03 | 0.00 | 77.13 | 0.80 | 78.33 | – |
| CTRL.Delta | – | – | – | 4.83 | -1.90 | 5.87 | 8.13 | -16.13 | -0.73 | -2.87 | 0.05 | 1.67 | – |

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 Above

Figure 7 shows the difference in the growth of percentages of students at math Proficient or Above, for the TRT and CTRL datasets, disaggregated by grade:

**Changes in Percent Proficient or Above – 2016/17 vs 2013/14**
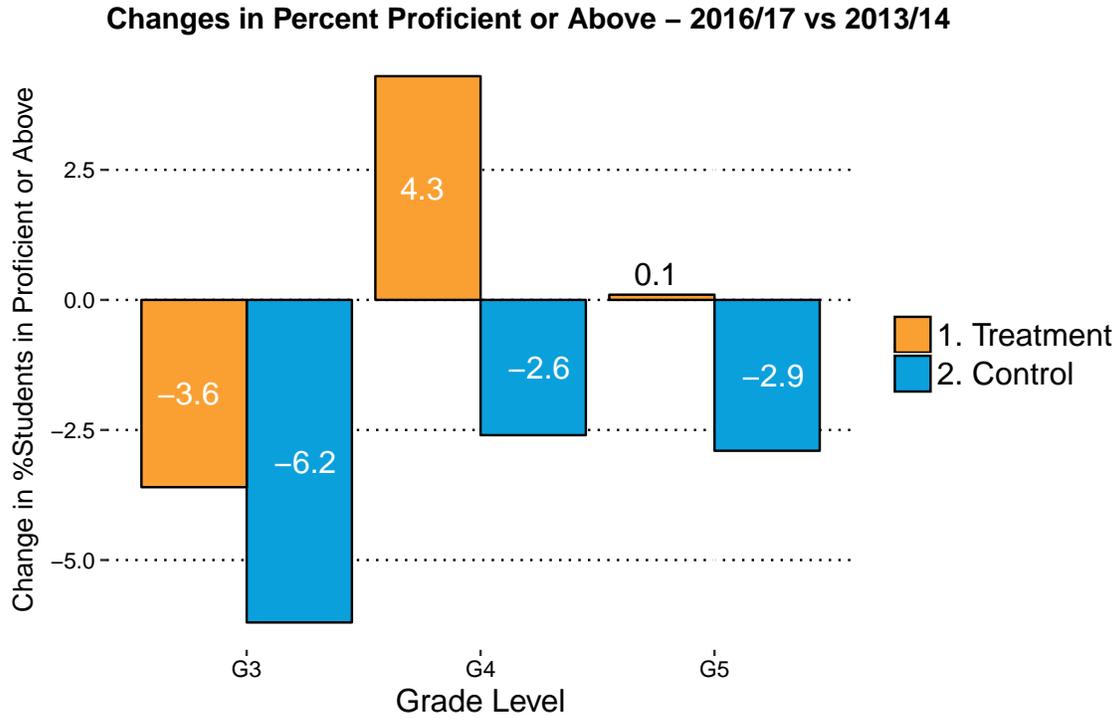


Figure 7: Changes in Percent of Students at Proficient or Above for TRT and CTRL Datasets between 2013/14 and 2016/17

Table 11 shows the statistics for the *differences* in changes between TRT and CTRL (Treatment - Control) for these same Proficient or Above math proficiency changes as shown in Figure 7.

|          | Estimate | P-Value | Int.Low | Int.High |
|----------|----------|---------|---------|----------|
| Grade 3  | 2.53     | 0.38    | -3.56   | 8.61     |
| Grade 4  | 6.94     | 0.02*   | 1.30    | 12.58    |

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

### 3.5.3 Grade-Level Analysis of Changes in OAA Z-scores

Figure 8 shows the changes in the grade-mean Z-scores of students for the TRT and CTRL datasets, disaggregated by grade:

**Changes in OAA Z–score – 2016/17 vs 2013/14**
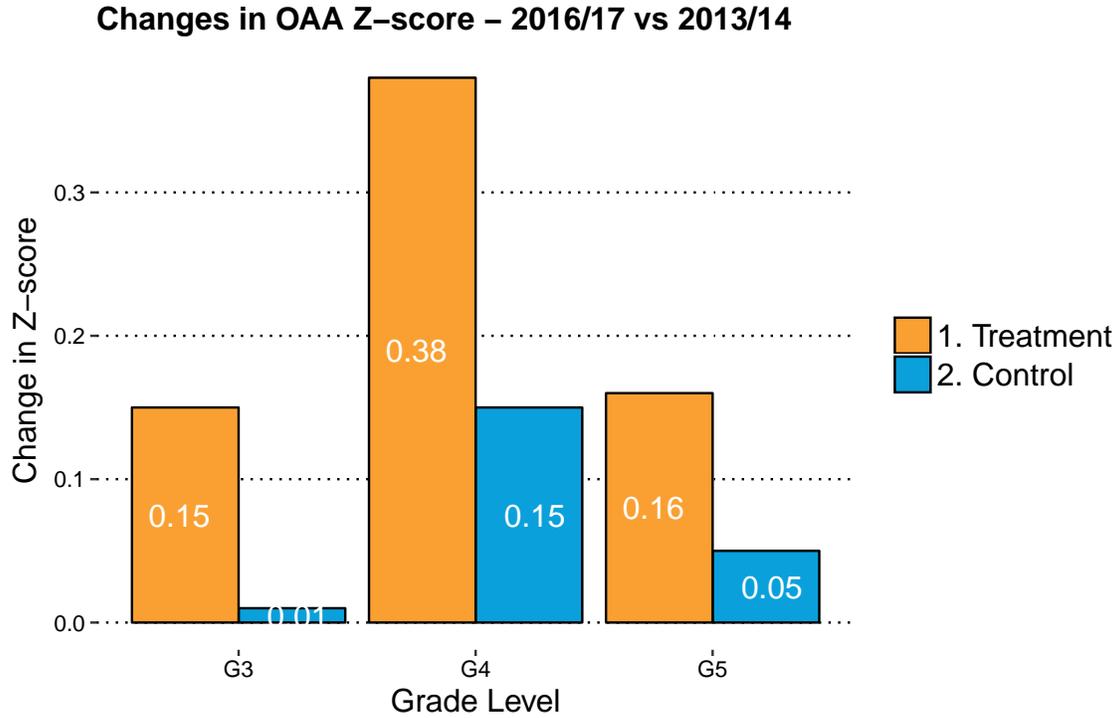


Figure 8: Changes in Grade-Mean OAA Z-score (See Section 3.1) for TRT and CTRL Datasets between 2013/14 and 2016/17

Table 12 shows the statistics for the differences between TRT and CTRL (Treatment - Control) for these same OAA z-score changes as shown in Figure 8.

|         | Estimate | P-Value | Int.Low | Int.High |
|---------|----------|---------|---------|----------|
| Grade 3 | 0.14     | 0.32    | -0.16   | 0.44     |
| Grade 4 | 0.23     | 0.07    | -0.02   | 0.49     |

Table 12: Statistics for the Differential Changes in OAA Z-scores (See Section 3.1) Growth, (TRT - CTRL)

# 4 Effect Size

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

|            | Proficient or Above Effect Size | Z-score Effect Size |
|------------|:-------------------------------:|:-------------------:|
| Grade 3    | 0.55                            | 0.52                |
| Grade 4    | 0.44                            | 0.29                |
| Grade 5    | 0.63                            | 0.55                |
| All Grades | 0.37                            | 0.29                |

Table 13: Cohen's d Effect Size

# 5 Findings Summary

Pickerington, Ohio grades 3, 4, and 5 using ST Math for the year 2016/17 averaged 65.7% ST Math Progress. 17/17 grades (100%) averaged covering more than 50% of ST Math content. Statistically significant differences were found in this analysis for both grade-aggregated and individual grade levels. Looking at Table 7, a statistically significant difference was found for grade-aggregated Proficient or Above proficiency levels, with a 4.42 point favorable differential for the ST Math treatment set. Further, in Table 7, grade-aggregated ST Math treatment set outperformed their matched controls at the Advanced level, with a statistically significant difference of 8.48. Referring to Table 11, a statistically significant difference was found for grade 4 Proficient or Above proficiency levels, with an estimate of 6.94 in favor of the ST Math treatment set.

# 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 14) and controls (Table 15) separately.

| | # Grades | # Schools | # Students | L1 | L2 | L3 | L4 | L5 | L6 | Proficient or Above | Z-score | Percentile | ST Math Per Comp. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Grade 3 (13.14) | 7 | 7 | 757 | 2.00 | 7.74 | 28.33 | 24.67 | 37.30 | 0.00 | 90.30 | 0.69 | 74.86 | – |
| Grade 4 (13.14) | 7 | 7 | 743 | 3.97 | 10.51 | 25.17 | 28.63 | 31.71 | 0.00 | 85.51 | 0.48 | 68.00 | – |
| Grade 5 (13.14) | 3 | 3 | 825 | 9.77 | 10.93 | 25.47 | 15.47 | 37.90 | 0.43 | 79.27 | 0.71 | 76.00 | – |
| All Grades (13.14) | 17 | 10 | 2325 | 4.18 | 9.45 | 26.52 | 24.68 | 35.11 | 0.08 | 86.38 | 0.61 | 72.24 | – |
| Grade 3 (14.15) | 7 | 7 | 680 | 2.84 | 10.73 | 23.13 | 49.33 | 13.97 | 0.00 | 86.43 | 0.98 | 83.00 | 56.45 |
| Grade 4 (14.15) | 7 | 7 | 725 | 2.07 | 11.19 | 25.07 | 54.83 | 6.81 | 0.00 | 86.71 | 1.08 | 85.71 | 50.75 |
| Grade 5 (14.15) | 3 | 3 | 726 | 1.60 | 15.20 | 30.53 | 44.30 | 8.20 | 0.17 | 83.20 | 0.89 | 80.67 | 50.54 |
| All Grades (14.15) | 17 | 10 | 2131 | 2.31 | 11.71 | 25.24 | 50.71 | 10.01 | 0.03 | 85.98 | 1.00 | 83.71 | 53.06 |
| Grade 3 (15.16) | 7 | 7 | 713 | 8.11 | 6.66 | 17.41 | 23.56 | 43.99 | 0.24 | 85.20 | 0.92 | 82.00 | 65.28 |
| Grade 4 (15.16) | 7 | 7 | 664 | 8.03 | 5.57 | 13.60 | 30.41 | 42.41 | 0.00 | 86.43 | 0.83 | 79.29 | 61.61 |
| Grade 5 (15.16) | 3 | 3 | 737 | 8.70 | 8.07 | 26.13 | 23.10 | 33.87 | 0.10 | 83.20 | 1.02 | 84.33 | 65.28 |
| All Grades (15.16) | 17 | 10 | 2114 | 8.18 | 6.46 | 17.38 | 26.30 | 41.55 | 0.12 | 85.35 | 0.90 | 81.29 | 63.77 |
| Grade 3 (16.17) | 7 | 7 | 728 | 7.49 | 5.89 | 17.56 | 27.64 | 41.47 | 0.00 | 86.67 | 0.84 | 79.29 | 64.06 |
| Grade 4 (16.17) | 7 | 7 | 776 | 6.64 | 3.44 | 14.37 | 29.63 | 45.53 | 0.31 | 89.84 | 0.86 | 80.29 | 67.19 |
| Grade 5 (16.17) | 3 | 3 | 741 | 12.53 | 8.13 | 25.73 | 23.50 | 30.10 | 0.00 | 79.33 | 0.88 | 80.67 | 66.02 |
| All Grades (16.17) | 17 | 10 | 2245 | 8.03 | 5.28 | 17.69 | 27.73 | 41.14 | 0.13 | 86.68 | 0.86 | 79.94 | 65.7 |

Table 14: TRT Grades Detail Sorted by Year

| | # Grades | # Schools | # Students | L1 | L2 | L3 | L4 | L5 | L6 | Proficient or Above | Z-score | Percentile | ST Math Per Comp. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Grade 3 (13.14) | 7 | 7 | 524 | 2.53 | 5.16 | 28.93 | 24.04 | 38.90 | 0.43 | 92.30 | 0.81 | 78.14 | – |
| Grade 4 (13.14) | 7 | 7 | 1095 | 6.91 | 12.00 | 29.39 | 24.97 | 26.73 | 0.00 | 81.09 | 0.25 | 61.14 | – |
| Grade 5 (13.14) | 3 | 3 | 624 | 9.27 | 10.63 | 23.63 | 17.47 | 38.17 | 0.73 | 80.00 | 0.75 | 76.67 | – |
| All Grades (13.14) | 17 | 17 | 2243 | 5.52 | 8.94 | 28.18 | 23.26 | 33.76 | 0.31 | 85.51 | 0.57 | 70.88 | – |
| Grade 3 (14.15) | 7 | 7 | 555 | 4.04 | 16.31 | 26.81 | 42.71 | 9.81 | 0.30 | 79.64 | 0.67 | 73.00 | – |
| Grade 4 (14.15) | 7 | 7 | 1052 | 16.11 | 16.16 | 28.07 | 37.67 | 2.00 | 0.00 | 67.74 | 0.25 | 61.71 | – |
| Grade 5 (14.15) | 3 | 3 | 510 | 3.97 | 18.43 | 32.30 | 39.77 | 5.57 | 0.00 | 77.63 | 0.66 | 72.67 | – |
| All Grades (14.15) | 17 | 17 | 2117 | 9.00 | 16.62 | 28.30 | 40.12 | 5.85 | 0.12 | 74.39 | 0.50 | 68.29 | – |
| Grade 3 (15.16) | 7 | 7 | 537 | 11.91 | 10.66 | 18.49 | 22.20 | 36.60 | 0.17 | 77.46 | 0.59 | 70.71 | – |
| Grade 4 (15.16) | 7 | 7 | 1066 | 15.64 | 7.59 | 16.27 | 33.21 | 27.29 | 0.00 | 76.77 | 0.45 | 67.86 | – |
| Grade 5 (15.16) | 3 | 3 | 586 | 10.10 | 11.33 | 26.57 | 25.13 | 26.83 | 0.00 | 78.53 | 0.84 | 79.00 | – |
| All Grades (15.16) | 17 | 17 | 2189 | 13.13 | 9.51 | 19.00 | 27.25 | 31.04 | 0.07 | 77.36 | 0.58 | 71.00 | – |
| Grade 3 (16.17) | 7 | 7 | 530 | 5.90 | 7.93 | 17.59 | 30.06 | 37.43 | 1.07 | 86.14 | 0.82 | 78.71 | – |
| Grade 4 (16.17) | 7 | 7 | 1064 | 14.44 | 7.07 | 20.66 | 28.17 | 29.17 | 0.47 | 78.47 | 0.40 | 65.14 | – |
| Grade 5 (16.17) | 3 | 3 | 620 | 14.10 | 8.73 | 29.50 | 25.60 | 22.03 | 0.00 | 77.13 | 0.80 | 78.33 | – |
| All Grades (16.17) | 17 | 17 | 2214 | 10.86 | 7.72 | 20.95 | 28.49 | 31.31 | 0.64 | 81.39 | 0.64 | 73.06 | – |

Table 15: 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 |
|---|---|---|---|---|
| 4917649 | DIL3BU | Pickerington Local | Diley Middle School | 5 |
| 2126131 | FAI3BU | Pickerington Local | Fairfield Elementary School | 3, 4 |
| 2042987 | HAR3BU | Pickerington Local | Harmon Middle School | 5 |
| 4938057 | HER3BU | Pickerington Local | Heritage Elementary School | 3, 4 |
| 799237 | PIC3BU | Pickerington Local | Pickerington Elementary School | 3, 4 |
| 11456523 | SYC3BU | Pickerington Local | Sycamore Creek Elementary School | 3, 4 |
| 11456509 | TOL3BV | Pickerington Local | Toll Gate Elementary | 3, 4 |
| 11456511 | TOL3BU | Pickerington Local | Toll Gate Middle School | 5 |
| 4449472 | TUS3BM | Pickerington Local | Tussing Elementary School | 3, 4 |
| 799251 | VIO3BU | Pickerington Local | Violet Elementary School | 3, 4 |

Table 16: 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 |
|---|---|---|---|
| 815635 | Amherst Exempted Village | Josiah Harris Elementary School | 4 |
| 5070658 | Avon Local | Avon Heritage Elementary School | 4 |
| 839978 | Bowling Green City | Conneaut Elementary School | 4 |
| 822236 | Buckeye Local | Buckeye Intermediate School | 4 |
| 806517 | Forest Hills Local | Maddux Elementary School | 3 |
| 814813 | Lakewood Local | Jackson Intermediate | 4 |
| 806799 | Lockland Local | Lockland Elementary School | 4 |
| 4747688 | Marysville Exempted Village | Mill Valley Elementary School | 4 |
| 813364 | Mentor Exempted Village | Bellflower Elementary School | 5 |
| 825446 | Northmont City | Northmoor Elementary School | 3 |
| 10002674 | Olentangy Local | Glen Oak Elementary School | 3 |
| 4281945 | Reynoldsburg City | Taylor Road Elementary School | 3 |
| 813572 | Riverside Local | Hale Road Elementary School | 3 |
| 817164 | Springfield Local | Dorr Street Elementary School | 3 |
| 4916920 | Wadsworth City | Central Intermediate School | 5 |
| 3392929 | West Branch Local | West Branch Middle School | 5 |
| 795035 | Westlake City | Holly Lane Elementary School | 3 |

Table 17: Matched Control Schools (CTRL Dataset)