

California Math Outcomes Analysis 2021/22

Grade Levels: 3, 4, 5
ST Math Program: Gen-6
Analysis Type: Z-score of scale score
Treatment-Years: 2021/22
Baseline-Year: 2018/19
Subgroup: All



Jessica Guise

© 2024-06-03

Abstract

This analysis evaluates grades using ST Math in California in 2021/22. 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 45 grades, consisting of grades 3, 4, and 5 at 30 schools, with an average baseline of 41% in Standard Met or Exceeded proficiency levels (refer to Figures 2 and 3 to see how your schools compare to those analyzed in this report). They were matched to 45 similar, randomly selected control grades at 44 schools that never used ST Math. Grade-wise growth in math proficiency was evaluated (i.e. growth in same grade, same school, from 2018/19 to 2021/22) 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 7.53 points at Standard Met or Exceeded, 1.42 points at Standard Met, 6.11 points at Standard Exceeded, and z-score of 0.36.

Contents

1	Introduction	5
1.1	Background	5
1.2	Program Description	5
2	Data Collection	6
2.1	Proficiency Levels Definition	6
2.2	Treatment Grades Pool and Selection	6
2.2.1	Enrollment Filter	6
2.2.2	Content Coverage Filter	6
2.3	Control Grades Pool and Selection	7
3	Data Analysis	8
3.1	Z-scores of Scale Score	8
3.2	Percentile Ranking	8
3.3	Final Treatment and Control	9
3.3.1	ST Math Grade-Aggregated Implementation ($\geq 85\%$ Enrollment Grades Only)	9
3.3.2	Filtering Treatment and Controls	10
3.3.3	Match of Controls to Treatment	11
3.4	Grade-Aggregated Analysis	13
3.5	Grade-Level Analysis	17
3.5.1	Grade Level Result Tables	17
3.5.2	Grade-Level Analysis of Changes in Math Standard Met or Exceeded	18
3.5.3	Grade-Level Analysis of Changes in CAASPP Math Scale Scores	19
3.5.4	Grade-Level Analysis of Changes in CAASPP Z-scores of Scale Scores	20
4	Effect Size	21
5	Findings Summary	21
6	Confounders	21
7	Reference Tables Grouped By School Year	22
8	Lists of Schools	23
8.1	Treatment Schools	23
8.2	Control Schools	24
9	Appendix	25

List of Figures

1	Histogram of ST Math Percent Progress for $\geq 85\%$ Enrollment Grades 2021/22	9
2	Baseline Year Density Plots Showing Math Scores Match between TRT and CTRL - 2018/19	11
3	Baseline Year Density Plot Showing Student Need Match between TRT and CTRL	12
4	Change at each Proficiency Level for Grade-Aggregated TRT and CTRL Datasets between 2018/19 and 2021/22	13
5	Changes in CAASPP Math scale scores and Z-scores (See Section 3.1) for Grade-Aggregated TRT and CTRL datasets between 2018/19 and 2021/22	14
6	Changes in Standard Met or Exceeded for Grade-Aggregated TRT and CTRL datasets between 2018/19 and 2021/22	15
7	Changes in Percentile Ranking for TRT and CTRL Datasets between 2018/19 and 2021/22	16
8	Changes in Percent of Students at Standard Met or Exceeded for TRT and CTRL Datasets between 2018/19 and 2021/22	18
9	Changes in Grade-Mean CAASPP Math scale score for TRT and CTRL Datasets between 2018/19 and 2021/22	19
10	Changes in Grade-Mean CAASPP Z-score of Scale Score (See Section 3.1) for TRT and CTRL Datasets between 2018/19 and 2021/22	20
11	Grade-aggregated Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	25
12	Grade-aggregated Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	26
13	Grade 3 Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	27
14	Grade 3 Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	28
15	Grade 4 Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	29
16	Grade 4 Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	30
17	Grade 5 Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	31
18	Grade 5 Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22	32

List of Tables

1	Proficiency Level Naming	6
2	Descriptive Statistics of ST Math Percent Progress for ≥ 85 percent Enrollment Grades	9
3	Number of ST Math Grades with ≥ 85 percent Enrollment and with ≥ 40 percent progress	9
4	Treatment Pool Filtering and Controls: Counts of Grades, Schools, and Students	10
5	Matching TRT and CTRL	12
6	Yearly Math Proficiency and Counts for TRT and CTRL Grade-Aggregated Datasets	13
7	Statistics for the Differential Changes in Math Scores Growth (TRT - CTRL)	15
8	Grade 3 - Yearly Math Performance and Counts for TRT and CTRL Datasets	17
9	Grade 4 - Yearly Math Performance and Counts for TRT and CTRL Datasets	17
10	Grade 5 - Yearly Math Performance and Counts for TRT and CTRL Datasets	17
11	Statistics for the Differential Changes in Standard Met or Exceeded, (TRT - CTRL)	18
12	Statistics for the Differential Changes in CAASPP Math scale scores Growth, (TRT - CTRL)	19
13	Statistics for the Differential Changes in CAASPP Z-scores of Scale Score (See Section 3.1) Growth, (TRT - CTRL)	20
14	Cohen's d Effect Size	21
15	TRT Grades Detail Sorted by Year	22
16	CTRL Grades Detail Sorted by Year	22
17	Treatment Schools (TRT Dataset)	23
18	Matched Control Schools (CTRL Dataset)	24
19	Statistics for the Grade-aggregated Match of Standard Met or Exceeded Between Treatment and Control	25
20	Statistics for the Match of Grade-aggregated Mean Scale Score Between Treatment and Control	26
21	Statistics for the Grade 3 Match of Standard Met or Exceeded Between Treatment and Control	27
22	Statistics for the Grade 3 Match of Mean Scale Score Between Treatment and Control	28
23	Statistics for the Grade 4 Match of Standard Met or Exceeded Between Treatment and Control	29
24	Statistics for the Grade 4 Match of Mean Scale Score Between Treatment and Control	30
25	Statistics for the Grade 5 Match of Standard Met or Exceeded Between Treatment and Control	31
26	Statistics for the Grade 5 Match of Mean Scale Score Between Treatment and Control	32

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 multi-year changes in grade-mean CAASPP Standard Met or Exceeded percentages. The treatment grades used the ST Math program for 2 or 3 years, beginning in the 2020/21 or 2019/20 school year, respectively. 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 (2018/19), and did not use ST Math in any subsequent year. The treatment grades’ selection pool was all schools using ST Math in grades 3, 4, and 5 in California. The control grades’ pool was all schools not using ST Math in grades 3, 4, and 5 in California. This study method measures effectiveness of the ST Math program when nominally implemented.

1.2 Program Description

Spatial-Temporal Math (ST Math) is game-based, instructional software for K–12 students, created by the MIND Research Institute (MIND). The purpose of the program is to boost math comprehension through visual learning. The ST Math software games begin without language or symbol abstractions by posing math problems as purely visual puzzles. In this way, three objectives 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. Interactive, animated visual manipulatives provide informative feedback on student solutions. A score of 100 percent on a game level comprised of 4-12 puzzles is required for progression through the levels. Failure requires a re-play of the level, via a new quasi-random set of puzzles. In this way, progression is self-paced.

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.

For students to achieve nominal progress through the program, there is a recommended time-on-task requirement of 90 minutes per week over about 30 weeks. Consistent application of 90 minutes per week throughout the school year is normally sufficient to result in a grade’s average ST Math content coverage exceeding 50% by year-end. In this study, we include grades that have achieved 40% or more content coverage (Progress) by April 15th.

This is a passive study with no experimental setup or extraordinary communications to any schools. All schools in this study therefore received normal program implementation support through the year from MIND support managers. This support includes bundled startup services of approximately 2-4 hours of training either in-person or online, access to live webinars, regular online and push reports on

usage and progress, email/phone helpdesk, and proactive monitoring for gaps or issues by MIND support representatives.

MIND Research Institute initiated, funded, and exercised editorial control over this study.

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 California's proficiency level descriptions:

Proficiency Level	State Proficiency Level Name
L1	Standard Not Met
L2	Standard Nearly Met
L3	Standard Met
L4	Standard Exceeded

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 California. From these schools, every grade that had used the ST Math program in 2021/22 was identified. They comprise the Treatment grades pool for this evaluation of multi-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 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. California's standardized math assessment (CAASPP). 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 California 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 40% by April.

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 40% 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 California. Though they are randomly selected, they are also matched to be similar to the Treatment grades' math attributes during the baseline year and the four years prior, in addition to the baseline demographics. The matched attributes include:

- scale score
- student percentages at Standard Met or Exceeded levels
- percentage of students receiving free or reduced lunch (using the demographic data from MDR).

The method of matching used is propensity score matching, via the "matchit" program in R, with "mahalanobis" as the distance measure. Refer to the Appendix for the full trend match of the baseline year and additional four prior years.

3 Data Analysis

The set of all schools and grades using ST Math in California is evaluated for Enrollment percentage and Progress percentage parameters. A filtered Treatment set (TRT) of all ST Math grades with $\geq 85\%$ Enrollment and $\geq 40\%$ 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 of Scale Score

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:

$$\begin{aligned} &\text{School A, Grade 3, Mean scale score: } 300 \\ &\text{Average across all schools statewide, Grade 3: } 350 \\ &\text{Standard deviation across all schools statewide, Grade 3: } 30 \\ &\text{Z-score} = ((\text{School A, Grade 3, Mean scale score}) - (\text{Average across all schools, Grade 3})) / (\text{Standard} \\ &\quad \text{deviation across all schools, Grade 3}) \\ &\text{Z-score} = \frac{300 - 350}{30} = -1.67 \end{aligned}$$

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

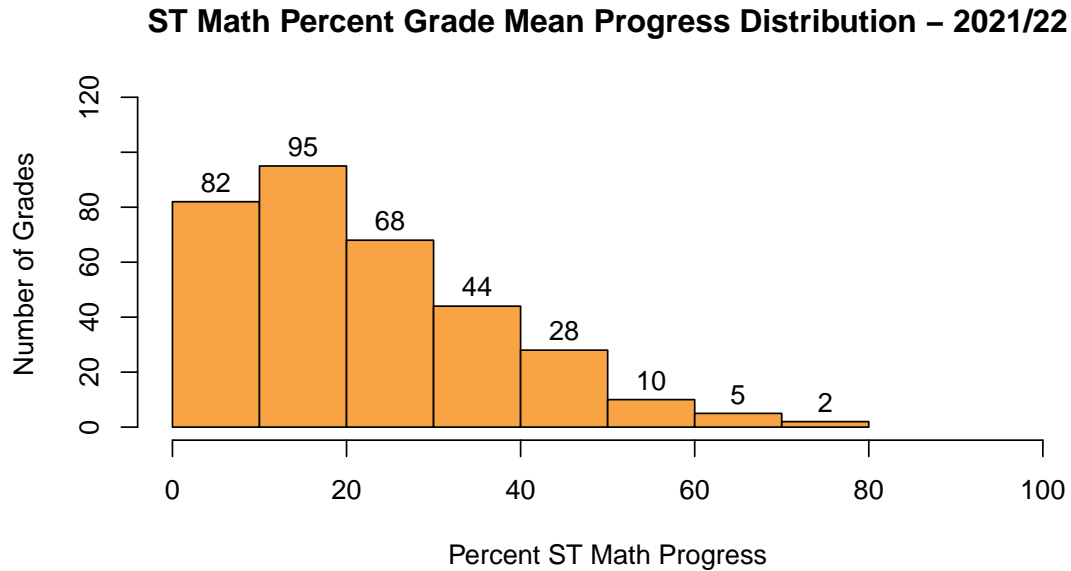


Figure 1: Histogram of ST Math Percent Progress for $\geq 85\%$ Enrollment Grades 2021/22

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 $\geq 40\%$ 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	0.0	99.6	12.5	14.0

Table 2: Descriptive Statistics of ST Math Percent Progress for $\geq 85\%$ Enrollment Grades

Grades with $\geq 85\%$ Enrollment:	334
Grades with in addition $\geq 40\%$ Progress:	45

Table 3: Number of ST Math Grades with $\geq 85\%$ Enrollment and with $\geq 40\%$ 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 filter, and also the 40% Progress filter. Other rows in the table indicate counts of numbers of students (2021/22 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	359	333	290	982
ST Math Using Schools	359	333	290	424
ST Math Students	25098	23523	21955	70576
ST Math Grades (Enroll \geq 85%)	118	112	104	334
TRT Grades (Enroll \geq 85% & Prog \geq 40%)	12	18	15	45
TRT Schools (Enroll \geq 85% & Prog \geq 40%)	12	18	15	30
TRT Students (Enroll \geq 85% & Prog \geq 40%)	1016	1413	1173	3602
CTRL Grades	12	18	15	45
CTRL Schools	12	17	15	44
CTRL Students	878	1278	1022	3178

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 CAASPP Math scale scores (left plot) and baseline percent students at CAASPP Standard Met or Exceeded (right plot) for treatment grades overlaid on control grades, showing the closeness of the match obtained between Treatment and Control sets of grades in the baseline year, 2018/19.

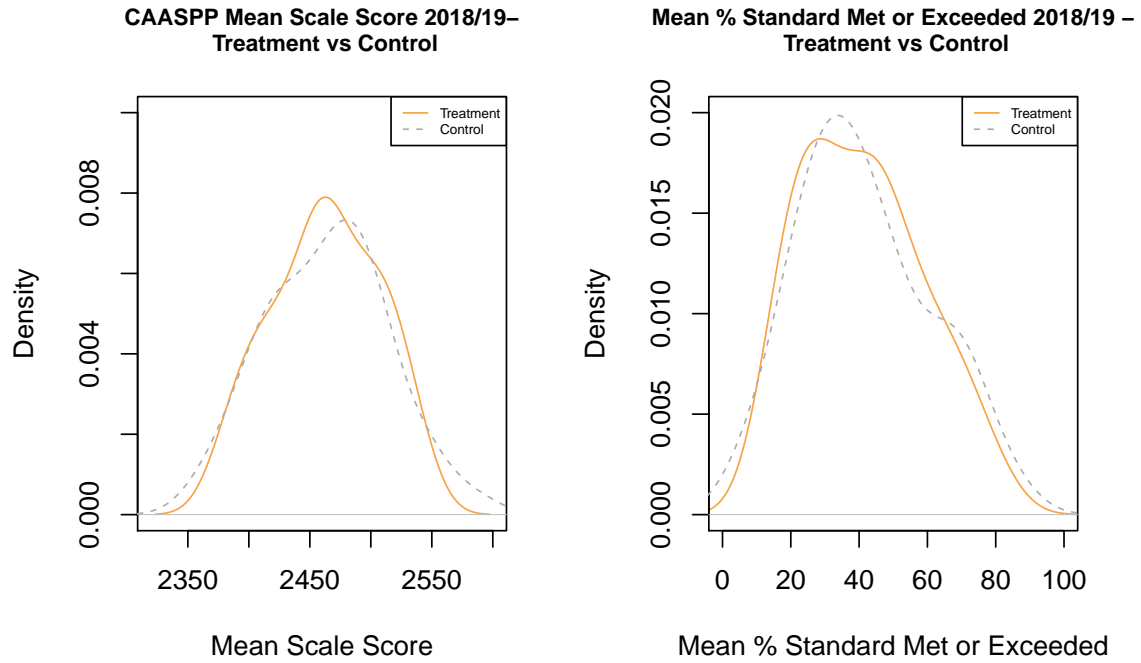


Figure 2: Baseline Year Density Plots Showing Math Scores Match between TRT and CTRL - 2018/19

Similarly, Figure 3 shows the density plot of the percentage of students needing free or reduced lunch for treatment grades overlaid on control grades, showing the closeness of the match obtained between Treatment and Control sets of grades.

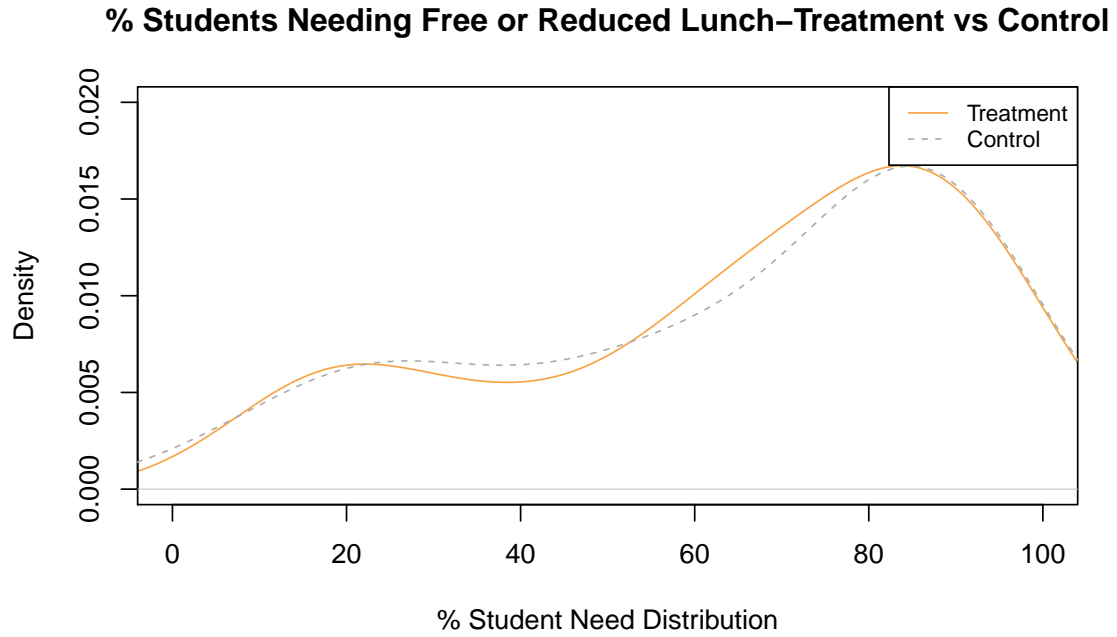


Figure 3: Baseline Year Density Plot Showing Student Need Match between TRT and CTRL

Table 5 shows the difference of the means of Treatment versus Control in the baseline year, with accompanying p-values, for percent Standard Met or Exceeded, 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	Effect Size
Standard Met or Exceeded - 2018/19	41.01	17.89	41.46	19.34	-0.45	0.91	-0.02
Scale Score - 2018/19	2462.93	44.04	2463.91	50.17	-0.98	0.92	-0.02
Percent Free or Reduced Lunch	65.33	26.60	64.49	27.51	0.84	0.88	0.03

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 distributions. The far right column also shows the average ST Math Progress for the TRT set.

	# Grades	# Schools	# Students	Scale Score	Z-Score of SS	Percentile	L1	L2	L3	L4	Standard Met or Exceeded	ST Math Per Comp.
TRT.18.19	45	30	3524	2462.9	-0.09	46.80	30.53	28.46	24.18	16.83	41.01	-
TRT.21.22	45	30	3477	2460.8	0.22	55.87	32.79	26.78	21.96	18.47	40.43	50.54
TRT.Delta	-	-	-	-2.2	0.31	9.07	2.26	-1.68	-2.23	1.64	-0.58	-
CTRL.18.19	45	44	3376	2463.9	-0.08	47.38	30.38	28.16	22.68	18.78	41.46	-
CTRL.21.22	45	44	3178	2443.3	-0.13	45.42	39.74	26.91	19.04	14.31	33.35	-
CTRL.Delta	-	-	-	-20.6	-0.05	-1.96	9.36	-1.25	-3.64	-4.47	-8.11	-

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).

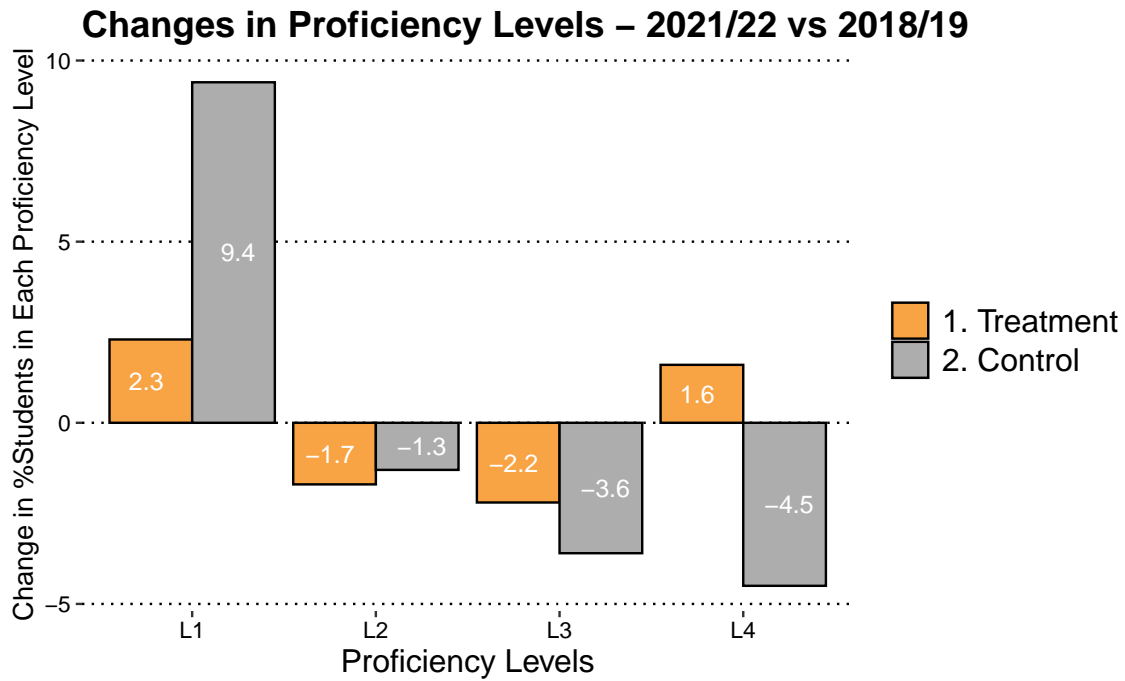


Figure 4: Change at each Proficiency Level for Grade-Aggregated TRT and CTRL Datasets between 2018/19 and 2021/22

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

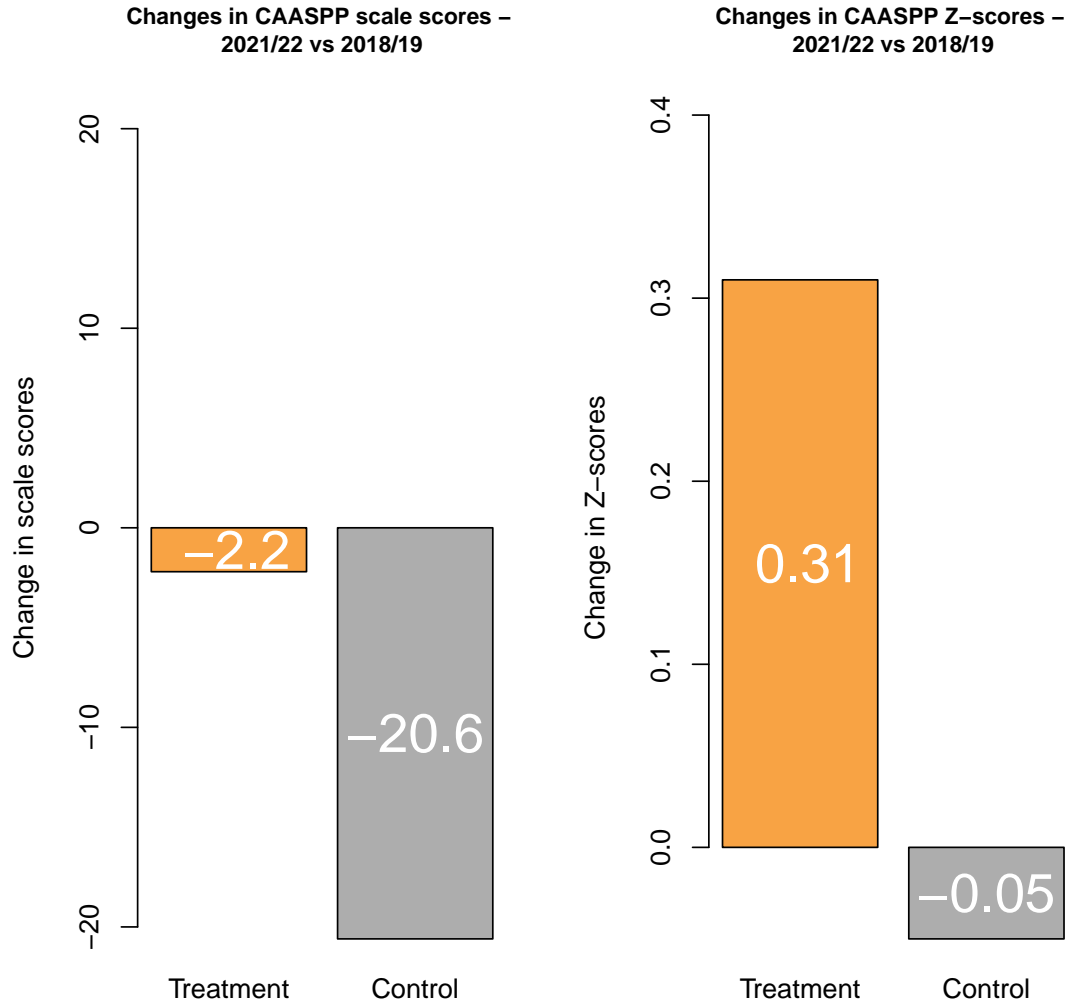


Figure 5: Changes in CAASPP Math scale scores and Z-scores (See Section 3.1) for Grade-Aggregated TRT and CTRL datasets between 2018/19 and 2021/22

Further, Figure 6 shows the changes in percent of students at CAASPP Standard Met or Exceeded for the grade-aggregated Treatment and Control sets.

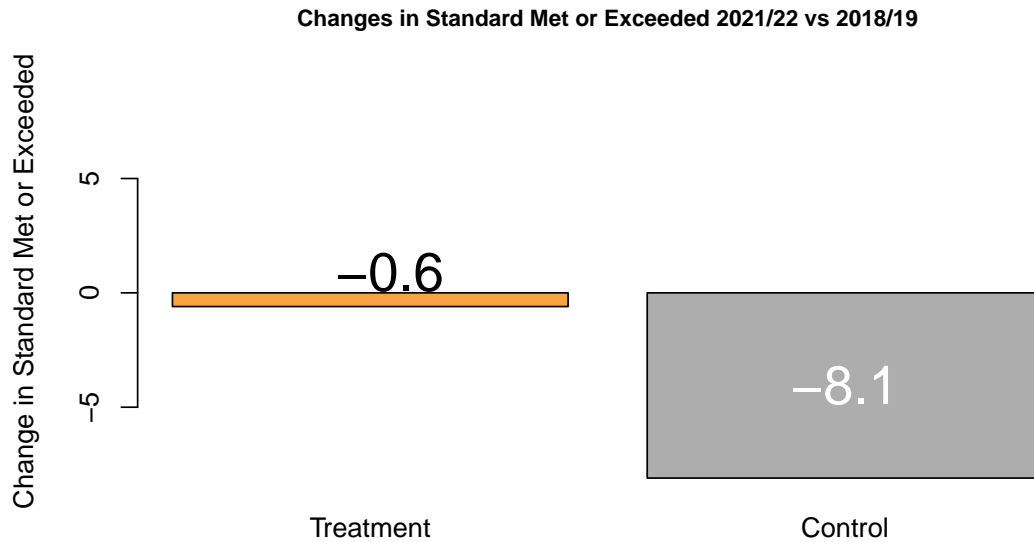


Figure 6: Changes in Standard Met or Exceeded for Grade-Aggregated TRT and CTRL datasets between 2018/19 and 2021/22

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

	Estimate	P-Value	Int.Low	Int.High
Standard Met or Exceeded	7.53	0.01*	2.22	12.84
Scale Score	18.41	0.00*	8.44	28.37
Z-score of SS	0.36	0.00*	0.15	0.57
L1	-7.11	0.01*	-12.03	-2.18
L2	-0.42	0.85	-5.02	4.17
L3	1.42	0.4	-1.90	4.74
L4	6.11	0.01*	1.87	10.35

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.

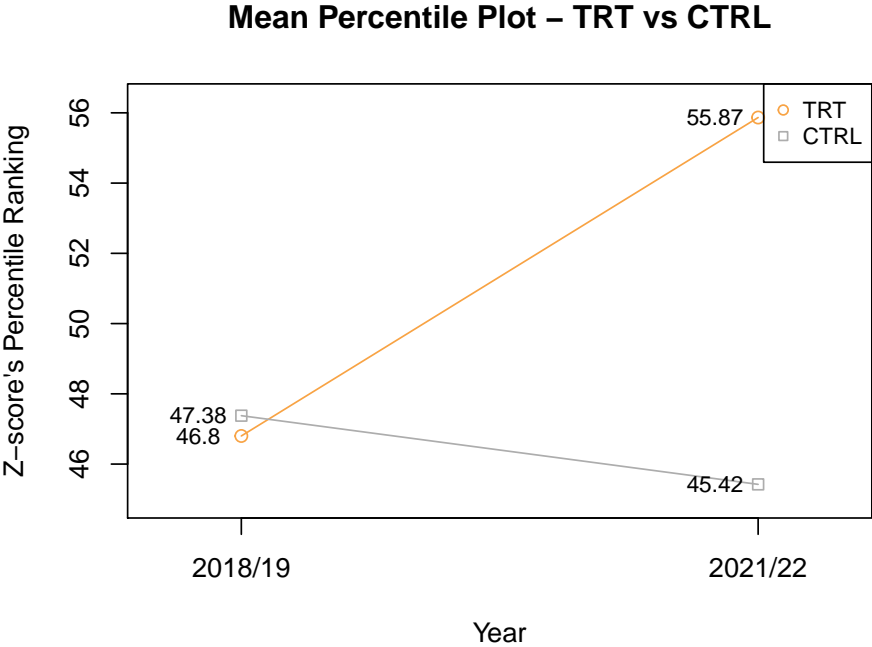


Figure 7: Changes in Percentile Ranking for TRT and CTRL Datasets between 2018/19 and 2021/22

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 of SS	Percentile	L1	L2	L3	L4	Standard Met or Exceeded	ST Math Per Prog.
TRT.18.19	12	12	1028	2425.6	-0.16	43.42	30.76	22.79	28.59	17.86	46.45	-
TRT.21.22	12	12	971	2425.2	0.20	55.17	32.95	19.30	26.84	20.92	47.75	53.11
TRT.Delta	-	-	-	-0.5	0.36	11.75	2.18	-3.49	-1.75	3.06	1.30	-
CTRL.18.19	12	12	961	2423.3	-0.21	41.50	31.46	23.73	26.30	18.51	44.81	-
CTRL.21.22	12	12	878	2401.6	-0.29	39.17	39.88	24.11	21.18	14.83	36.01	-
CTRL.Delta	-	-	-	-21.7	-0.07	-2.33	8.42	0.39	-5.12	-3.69	-8.80	-

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

	# Grades	# Schools	# Students	Scale Score	Z-score of SS	Percentile	L1	L2	L3	L4	Standard Met or Exceeded	ST Math Per Prog.
TRT.18.19	18	18	1359	2470.2	0.01	51.22	23.87	32.81	26.10	17.23	43.33	-
TRT.21.22	18	18	1387	2466.8	0.28	57.83	27.67	31.22	22.10	19.01	41.11	50.12
TRT.Delta	-	-	-	-3.4	0.27	6.61	3.80	-1.58	-4.00	1.78	-2.21	-
CTRL.18.19	18	17	1301	2464.2	-0.13	47.94	26.73	31.84	23.93	17.49	41.42	-
CTRL.21.22	18	17	1278	2440.2	-0.24	44.11	38.70	28.34	21.51	11.45	32.96	-
CTRL.Delta	-	-	-	-24.0	-0.12	-3.83	11.97	-3.50	-2.42	-6.04	-8.46	-

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

	# Grades	# Schools	# Students	Scale Score	Z-score of SS	Percentile	L1	L2	L3	L4	Standard Met or Exceeded	ST Math Per Prog.
TRT.18.19	15	15	1137	2484.0	-0.17	44.20	38.34	27.78	18.36	15.53	33.88	-
TRT.21.22	15	15	1119	2482.0	0.15	54.07	38.81	27.44	17.88	15.86	33.75	48.98
TRT.Delta	-	-	-	-2.1	0.32	9.87	0.47	-0.34	-0.47	0.34	-0.13	-
CTRL.18.19	15	15	1114	2496.0	0.07	51.40	33.89	27.28	18.28	20.54	38.83	-
CTRL.21.22	15	15	1022	2480.6	0.13	52.00	40.88	27.42	14.36	17.33	31.69	-
CTRL.Delta	-	-	-	-15.5	0.05	0.60	7.00	0.14	-3.93	-3.21	-7.14	-

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

3.5.2 Grade-Level Analysis of Changes in Math Standard Met or Exceeded

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

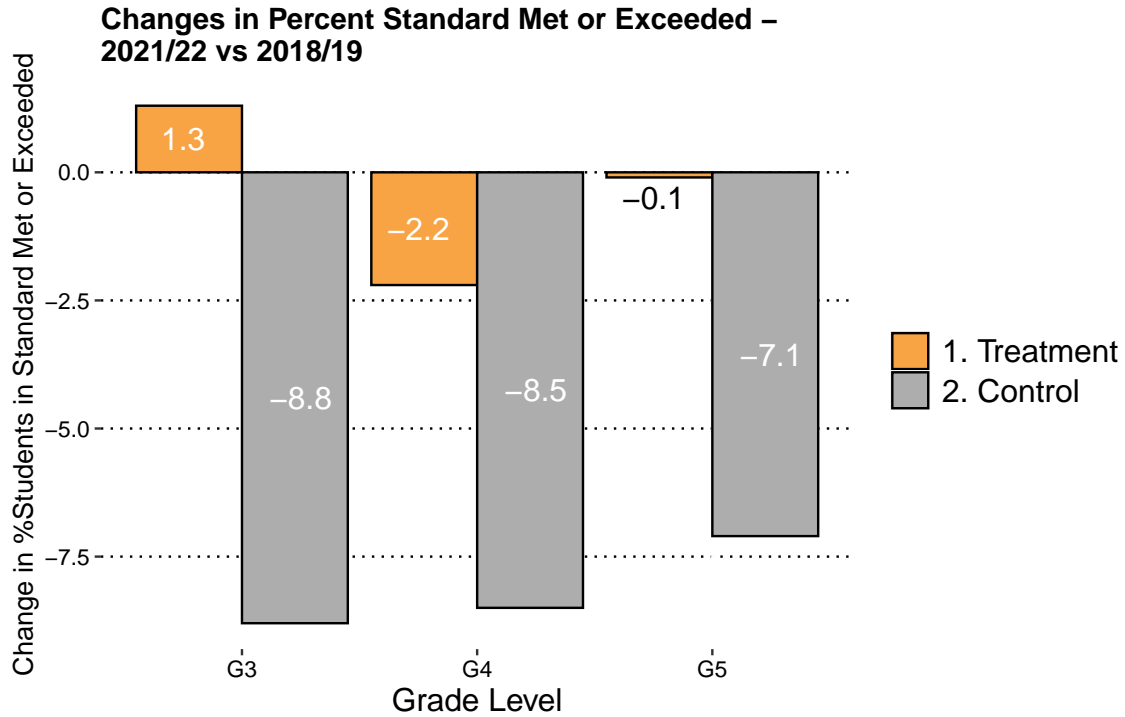


Figure 8: Changes in Percent of Students at Standard Met or Exceeded for TRT and CTRL Datasets between 2018/19 and 2021/22

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

	Estimate	P-Value	Int.Low	Int.High
Grade 3	10.11	0.15	-4.05	24.27
Grade 4	6.25	0.14	-2.14	14.63
Grade 5	7.00	0.07	-0.63	14.63

Table 11: Statistics for the Differential Changes in Standard Met or Exceeded, (TRT - CTRL)

3.5.3 Grade-Level Analysis of Changes in CAASPP 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:

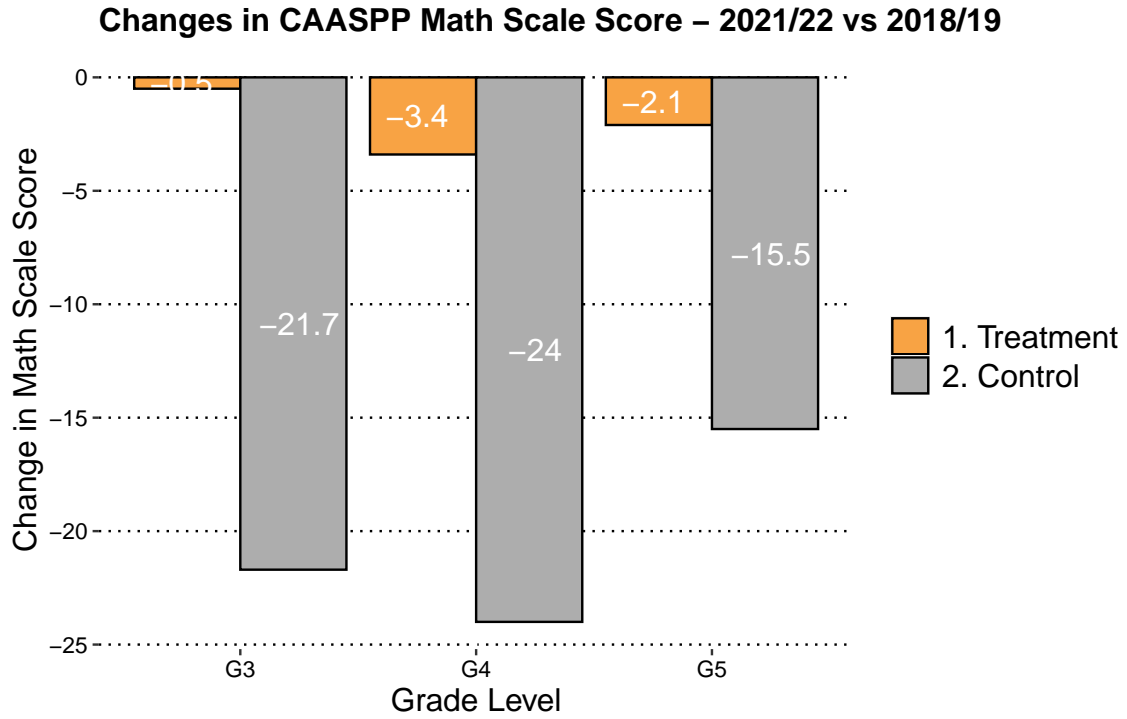


Figure 9: Changes in Grade-Mean CAASPP Math scale score for TRT and CTRL Datasets between 2018/19 and 2021/22

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

	Estimate	P-Value	Int.Low	Int.High
Grade 3	21.28	0.13	-6.69	49.24
Grade 4	20.66	0.01*	6.49	34.84
Grade 5	13.41	0.07	-1.33	28.15

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

3.5.4 Grade-Level Analysis of Changes in CAASPP Z-scores of Scale Scores

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

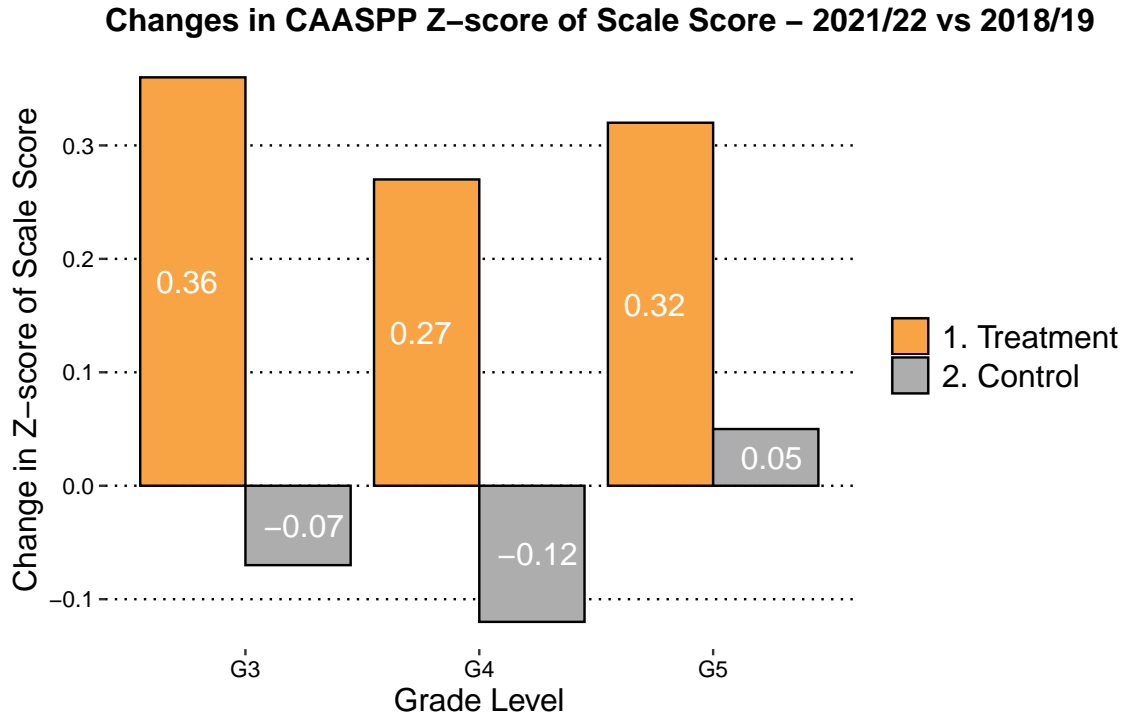


Figure 10: Changes in Grade-Mean CAASPP Z-score of Scale Score (See Section 3.1) for TRT and CTRL Datasets between 2018/19 and 2021/22

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

	Estimate	P-Value	Int.Low	Int.High
Grade 3	0.43	0.16	-0.18	1.04
Grade 4	0.39	0.01*	0.09	0.69
Grade 5	0.27	0.06	-0.01	0.55

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

4 Effect Size

The following table shows the effect sizes for Standard Met or Exceeded, CAASPP scale score, and accompanying Z-score.

	Scale score Effect Size	Z-score of Scale Score Effect Size	Standard Met or Exceeded Effect Size
Grade 3	0.54	0.45	0.50
Grade 4	0.49	0.41	0.32
Grade 5	0.30	0.29	0.36
All Grades	0.39	0.39	0.40

Table 14: Cohen's d Effect Size

5 Findings Summary

California grades 3, 4, and 5 using ST Math for the year 2021/22 averaged 12.5% ST Math Progress. 56/982 grades (6%) averaged covering more than 40% of ST Math content. Statistically significant differences were found in this analysis for both grade-aggregated and individual grade levels. Looking at Table 7, statistically significant differences were found for grade-aggregated z-score of scale score, with an estimate of 0.36 points favorable for the ST Math treatment set, as well as for grade-aggregated Standard Met or Exceeded proficiency levels, with a 7.53 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 Standard Exceeded level, with a statistically significant difference of 6.11. Looking at Table 13, grade 4 ST math treatment set outperformed their matched controls for CAASPP z-score of scale score with a statistically significant difference of 0.39.

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 40% 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.

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 of SS	Percentile	L1	L2	L3	L4	Standard Met or Exceeded	ST Math Per Comp.
Grade 3 (18.19)	12	12	1028	2425.6	-0.16	43.42	30.76	22.79	28.59	17.86	46.45	-
Grade 4 (18.19)	18	18	1359	2470.2	0.01	51.22	23.87	32.81	26.10	17.23	43.33	-
Grade 5 (18.19)	15	15	1137	2484.0	-0.17	44.20	38.34	27.78	18.36	15.53	33.88	-
All Grades (18.19)	45	30	3524	2462.9	-0.09	46.80	30.53	28.46	24.18	16.83	41.01	-
Grade 3 (21.22)	12	12	971	2425.2	0.20	55.17	32.95	19.30	26.84	20.92	47.75	53.11
Grade 4 (21.22)	18	18	1387	2466.8	0.28	57.83	27.67	31.22	22.10	19.01	41.11	50.12
Grade 5 (21.22)	15	15	1119	2482.0	0.15	54.07	38.81	27.44	17.88	15.86	33.75	48.98
All Grades (21.22)	45	30	3477	2460.8	0.22	55.87	32.79	26.78	21.96	18.47	40.43	50.54

Table 15: TRT Grades Detail Sorted by Year

	# Grades	# Schools	# Students	Scale Score	Z-Score of SS	Percentile	L1	L2	L3	L4	Standard Met or Exceeded	ST Math Per Comp.
Grade 3 (18.19)	12	12	961	2423.3	-0.21	41.50	31.46	23.73	26.30	18.51	44.81	-
Grade 4 (18.19)	18	17	1301	2464.2	-0.13	47.94	26.73	31.84	23.93	17.49	41.42	-
Grade 5 (18.19)	15	15	1114	2496.0	0.07	51.40	33.89	27.28	18.28	20.54	38.83	-
All Grades (18.19)	45	44	3376	2463.9	-0.08	47.38	30.38	28.16	22.68	18.78	41.46	-
Grade 3 (21.22)	12	12	878	2401.6	-0.29	39.17	39.88	24.11	21.18	14.83	36.01	-
Grade 4 (21.22)	18	17	1278	2440.2	-0.24	44.11	38.70	28.34	21.51	11.45	32.96	-
Grade 5 (21.22)	15	15	1022	2480.6	0.13	52.00	40.88	27.42	14.36	17.33	31.69	-
All Grades (21.22)	45	44	3178	2443.3	-0.13	45.42	39.74	26.91	19.04	14.31	33.35	-

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 40% progress filtering) used in the analysis.

PID	District	School Name	GRADE
4455342	Bellflower Unified	Albert Baxter Elementary	3, 5
66717	Bellflower Unified	Stephen Foster Elementary	4
68820	El Monte City	Columbia Elementary	4, 5
68844	El Monte City	Durfee Elementary	3, 4
68856	El Monte City	Gidley Elementary	4
137990	Hope Elementary	Hope Elementary	4, 5
91011	Le Grand Union Elementary	Le Grand Elementary	5
76554	Los Angeles Unified	Brainard Elementary	3
76475	Los Angeles Unified	Rockdale Visual & Performing Arts Magnet	3, 4, 5
11829889	Los Angeles Unified	Sylmar Leadership Academy	5
91255	Merced City Elementary	John C. Fremont Elementary	3
3054830	Oxnard	Christa McAuliffe Elementary	3, 4
140040	Oxnard	Curren Elementary	3, 4
140052	Oxnard	Driffill Elementary	3
4017110	Oxnard	Emilie Ritche Elementary	5
140076	Oxnard	Harrington Elementary	5
11447077	Oxnard	Juan Lagunas Soria Elementary	4
140117	Oxnard	Kamala Elementary	4, 5
140129	Oxnard	Marina West Elementary	5
140131	Oxnard	McKinna Elementary	4
4748682	Oxnard	Norman R. Brekke Elementary	3, 4, 5
140167	Oxnard	Sierra Linda Elementary	3
2104119	Rocklin Unified	Rocklin Elementary	4, 5
4304171	Saddleback Valley Unified	Foothill Ranch Elementary	3
94180	Saddleback Valley Unified	Montevideo Elementary	4
4289234	Saddleback Valley Unified	Robinson Elementary	4, 5
3011905	Saddleback Valley Unified	Trabuco Mesa Elementary	3, 4, 5
4755142	San Marcos Unified	Discovery Elementary	4
4032495	San Mateo-Foster City	Fiesta Gardens International Elementary	5
11829920	Sierra Foothill Charter	Sierra Foothill Charter	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
65634	ABC Unified	Furgeson (Venn W.) Elementary	3
102892	Arcohe Union Elementary	Arcohe Elementary	5
66913	Burbank Unified	Thomas Jefferson Elementary	3
67498	Claremont Unified	Mountain View Elementary	5
67826	Compton Unified	Ralph Bunche Elementary	4
101379	Corona-Norco Unified	Vicentia Elementary	5
4456425	Dos Palos Oro Loma Joint Unified	Bernhard Marks Elementary	3
4916164	Etiwanda Elementary	Cecilia Lucero Solorio Elementary	4
107842	Fontana Unified	Virginia Primrose Elementary	3
57144	Fresno Unified	Centennial Elementary	3
96645	Garden Grove Unified	Stanford Elementary	4
70093	Hacienda la Puente Unified	Grandview, College Preparatory Academy	4
4375572	Hart-Ransom Union Elementary	Hart-Ransom Academic Charter	4
48997	Livermore Valley Joint Unified	Joe Michell	4
3316482	Livingston Union	Yamato Colony Elementary	4
120909	Lucia Mar Unified	Ocean View Elementary	4
88571	Madera Unified	Alpha Elementary	3
10908004	Madera Unified	John J. Pershing Elementary	5
135588, 3318923	Modesto City Elementary, Martinez Unified	John Muir Elementary	4, 4
3253828	Moreno Valley Unified	Hendrick Ranch Elementary	3
4871966	Oak Grove Elementary	Ledesma (Rita) Elementary	5
12032534	Oakley Union Elementary	Almond Grove Elementary	4
108224	Ontario-Montclair	Bon View Elementary	4
108456	Ontario-Montclair	Moreno Elementary	5
102189	Palm Springs Unified	Vista del Monte Elementary	5
128078	Palo Alto Unified	Walter Hays Elementary	3
2846240	Panama-Buena Vista Union	Laurelglen Elementary	5
5119008	Pathways Charter	Pathways Charter	3
102282	Perris Elementary	Perris Elementary	4
98423	Placentia-Yorba Linda Unified	Glenknoll Elementary	5
1169946	Redlands Unified	Mariposa Elementary	4
64769	Reef-Sunset Unified	Kettleman City Elementary	3
102567	Riverside Unified	Madison Elementary	5
4020014	Rocklin Unified	Antelope Creek Elementary	3
104670	Sacramento City Unified	John H. Still	5
116611	San Francisco Unified	Chavez (Cesar) Elementary	4
116788	San Francisco Unified	Serra (Junipero) Elementary	3
50859	San Lorenzo Unified	Corvallis Elementary	4
4747145	San Ramon Valley Unified	Tassajara Hills Elementary	5
109606	Upland Unified	Citrus Elementary	5
4914219	Victor Elementary	Mountain View Montessori Charter	5
55108	West Contra Costa Unified	Mira Vista Elementary	4
55251	West Contra Costa Unified	Stege Elementary	4
141991	Woodland Joint Unified	Rhoda Maxwell Elementary	5

Table 18: Matched Control Schools (CTRL Dataset)

9 Appendix

Figure 11 charts the grade-aggregated trends of treatment and control for percent of students at Standard Met or Exceeded.

Mean % Standard Met or Exceeded Plot–Treatment vs Control

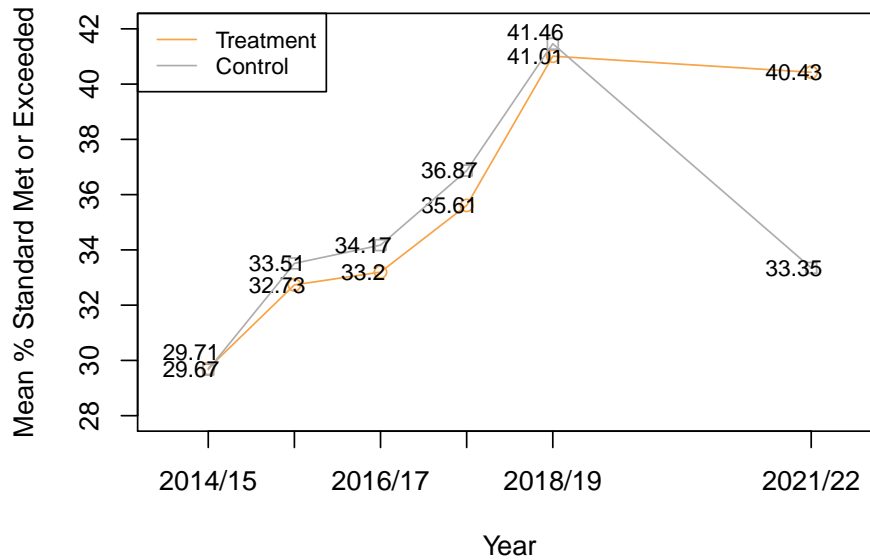


Figure 11: Grade-aggregated Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 19 shows the statistics for the grade-aggregated match between Treatment and Control for these same Standard Met or Exceeded changes as shown in Figure 11.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	29.71	29.67	0.99	0.00
2015/16	32.73	33.51	0.86	-0.04
2016/17	33.20	34.17	0.82	-0.05
2017/18	35.61	36.87	0.77	-0.06
2018/19	41.01	41.46	0.91	-0.02

Table 19: Statistics for the Grade-aggregated Match of Standard Met or Exceeded Between Treatment and Control

Figure 12 charts the grade-aggregated trends of treatment and control for mean scale score.

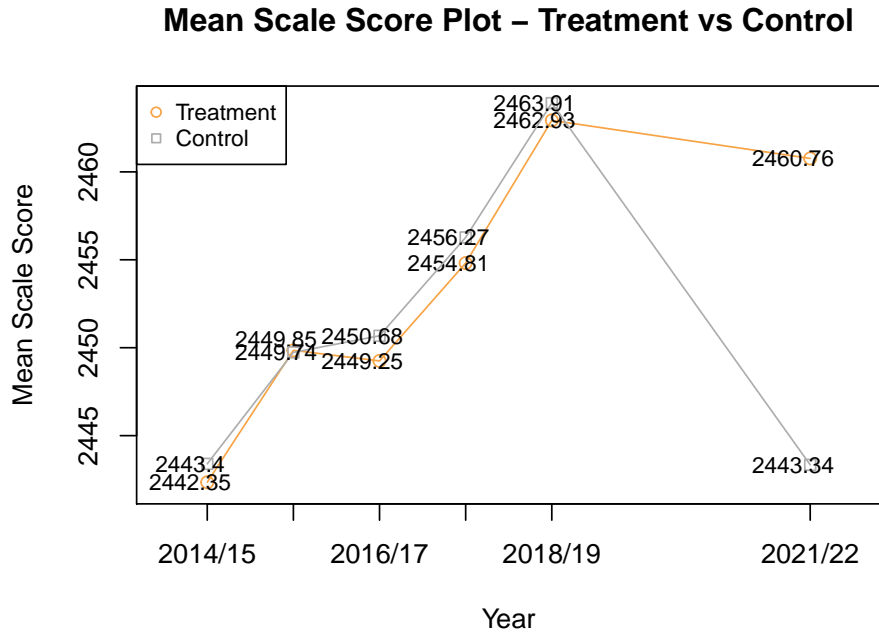


Figure 12: Grade-aggregated Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 20 shows the statistics for the grade-aggregated match between Treatment and Control for these same mean scale score changes as shown in Figure 12.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	2442.35	2443.40	0.92	-0.02
2015/16	2449.85	2449.74	0.99	0.00
2016/17	2449.25	2450.68	0.89	-0.03
2017/18	2454.81	2456.27	0.89	-0.03
2018/19	2462.93	2463.91	0.92	-0.02

Table 20: Statistics for the Match of Grade-aggregated Mean Scale Score Between Treatment and Control

Figure 13 charts the grade 3 trends of treatment and control for percent of students at Standard Met or Exceeded.

Mean % Standard Met or Exceeded Plot – Treatment vs Control

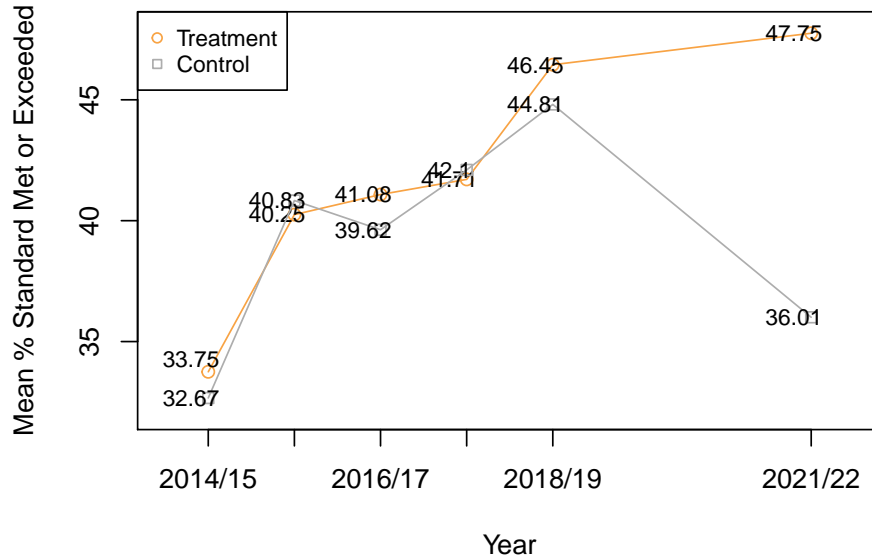


Figure 13: Grade 3 Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 21 shows the statistics for the grade 3 match between Treatment and Control for these same Standard Met or Exceeded changes as shown in Figure 13.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	33.75	32.67	0.90	0.05
2015/16	40.25	40.83	0.95	-0.02
2016/17	41.08	39.62	0.86	0.07
2017/18	41.71	42.10	0.97	-0.02
2018/19	46.45	44.81	0.84	0.08

Table 21: Statistics for the Grade 3 Match of Standard Met or Exceeded Between Treatment and Control

Figure 14 charts the grade 3 trends of treatment and control for mean scale score.

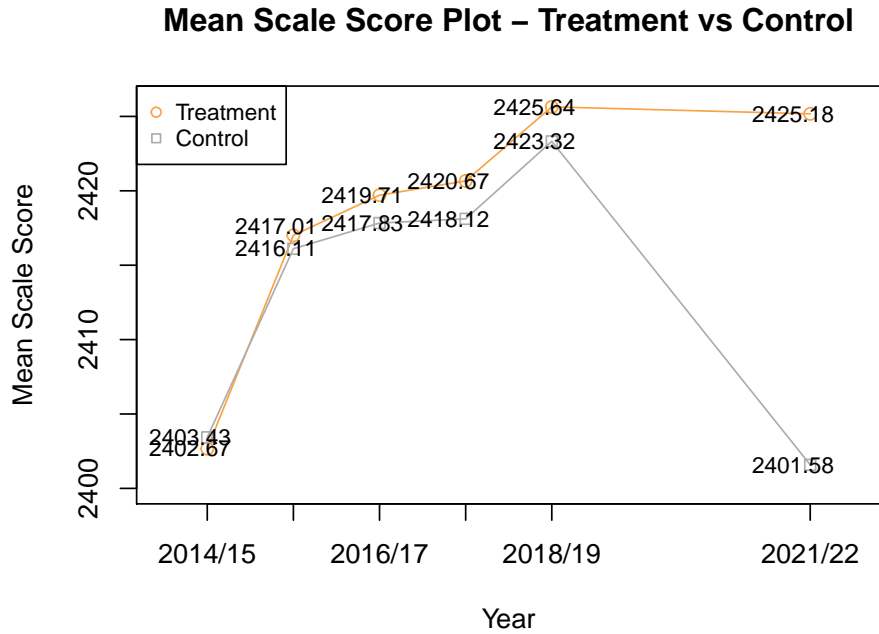


Figure 14: Grade 3 Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 22 shows the statistics for the grade 3 match between Treatment and Control for these same mean scale score changes as shown in Figure 14.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	2402.67	2403.43	0.96	-0.02
2015/16	2417.01	2416.11	0.96	0.02
2016/17	2419.71	2417.83	0.91	0.05
2017/18	2420.67	2418.12	0.89	0.06
2018/19	2425.64	2423.32	0.88	0.06

Table 22: Statistics for the Grade 3 Match of Mean Scale Score Between Treatment and Control

Figure 15 charts the grade 4 trends of treatment and control for percent of students at Standard Met or Exceeded.

Mean % Standard Met or Exceeded Plot – Treatment vs Control

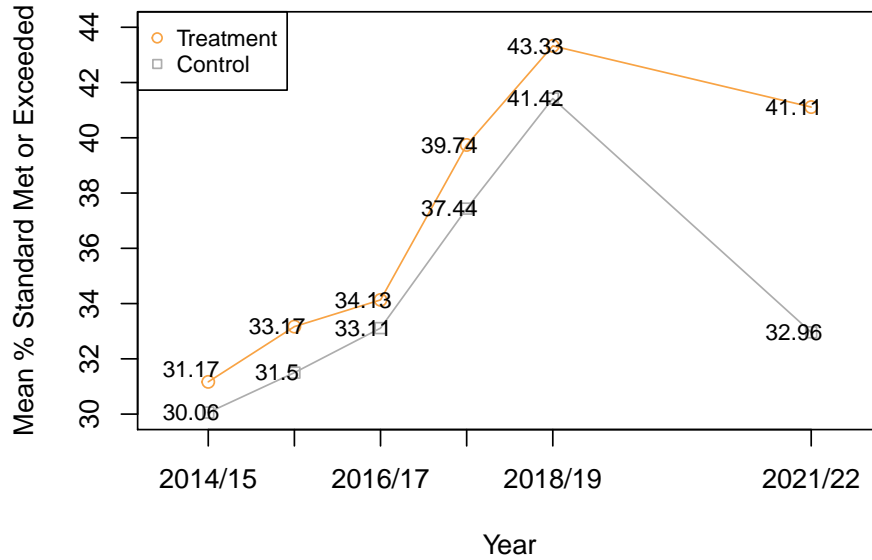


Figure 15: Grade 4 Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 23 shows the statistics for the grade 4 match between Treatment and Control for these same Standard Met or Exceeded changes as shown in Figure 15.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	31.17	30.06	0.88	0.05
2015/16	33.17	31.50	0.80	0.08
2016/17	34.13	33.11	0.88	0.05
2017/18	39.74	37.44	0.73	0.11
2018/19	43.33	41.42	0.76	0.10

Table 23: Statistics for the Grade 4 Match of Standard Met or Exceeded Between Treatment and Control

Figure 16 charts the grade 4 trends of treatment and control for mean scale score.

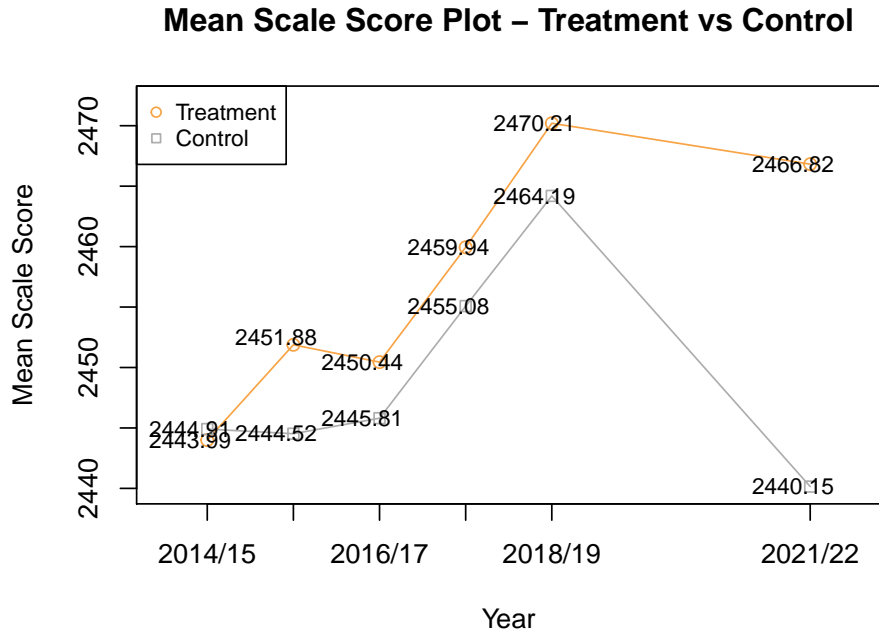


Figure 16: Grade 4 Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 24 shows the statistics for the grade 4 match between Treatment and Control for these same mean scale score changes as shown in Figure 16.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	2443.99	2444.91	0.95	-0.02
2015/16	2451.88	2444.52	0.61	0.17
2016/17	2450.44	2445.81	0.74	0.11
2017/18	2459.94	2455.08	0.73	0.11
2018/19	2470.21	2464.19	0.67	0.14

Table 24: Statistics for the Grade 4 Match of Mean Scale Score Between Treatment and Control

Figure 17 charts the grade 5 trends of treatment and control for percent of students at Standard Met or Exceeded.

Mean % Standard Met or Exceeded Plot – Treatment vs Control

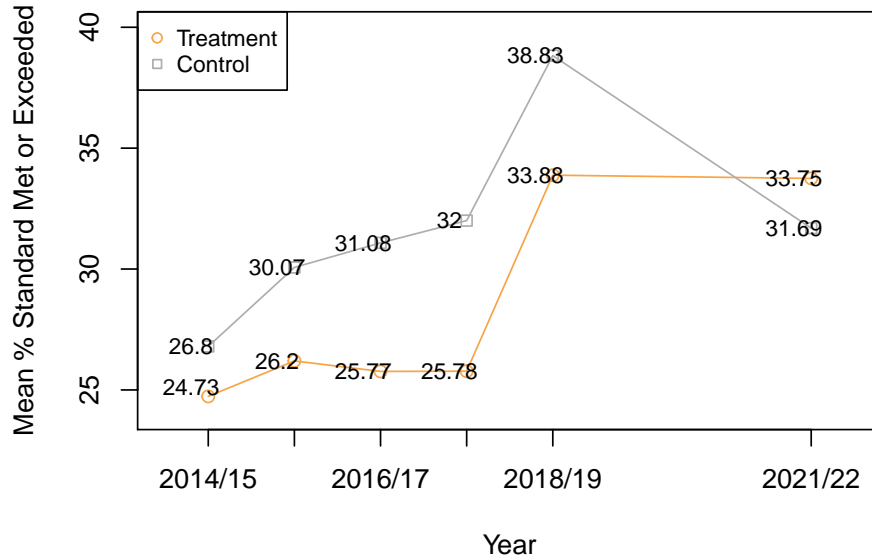


Figure 17: Grade 5 Match of Standard Met or Exceeded for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 25 shows the statistics for the grade 5 match between Treatment and Control for these same Standard Met or Exceeded changes as shown in Figure 17.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	24.73	26.80	0.77	-0.10
2015/16	26.20	30.07	0.62	-0.18
2016/17	25.77	31.08	0.49	-0.25
2017/18	25.78	32.00	0.38	-0.32
2018/19	33.88	38.83	0.44	-0.28

Table 25: Statistics for the Grade 5 Match of Standard Met or Exceeded Between Treatment and Control

Figure 18 charts the grade 5 trends of treatment and control for mean scale score.

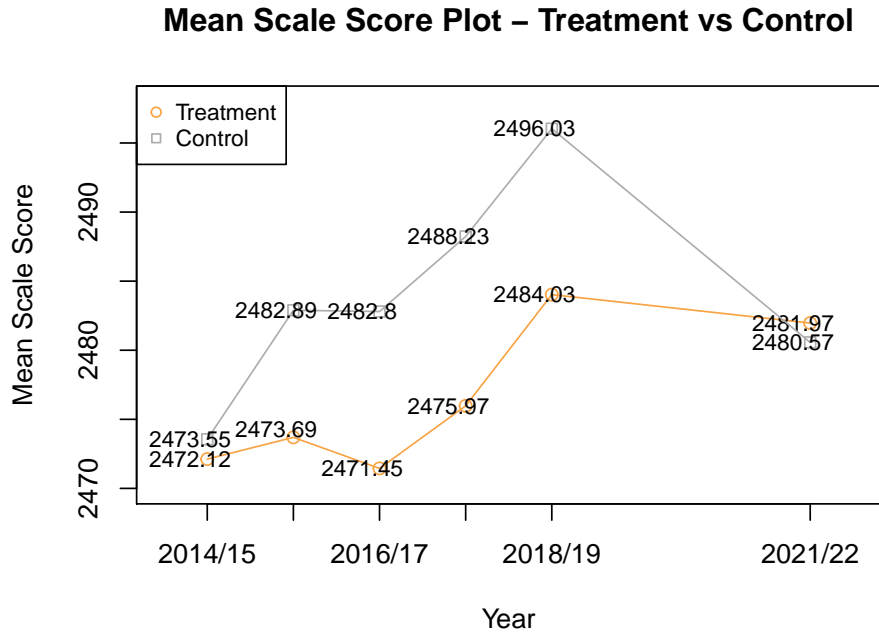


Figure 18: Grade 5 Match of Mean Scale Score for Treatment and Control Datasets for 2014/15, 2015/16, 2016/17, 2017/18, 2018/19, in addition to changes between 2018/19 and 2021/22

Table 26 shows the statistics for the grade 5 match between Treatment and Control for these same mean scale score changes as shown in Figure 18.

	TRT	CTRL	P-Value	Effect Size (Hedges' G)
2014/15	2472.12	2473.55	0.93	-0.03
2015/16	2473.69	2482.89	0.60	-0.19
2016/17	2471.45	2482.80	0.55	-0.22
2017/18	2475.97	2488.23	0.46	-0.27
2018/19	2484.03	2496.03	0.42	-0.29

Table 26: Statistics for the Grade 5 Match of Mean Scale Score Between Treatment and Control