2020

K-12 Mathematics Personalized Learning Program 2018-19 Student Outcomes, Evaluation Addendum for Year 3

Prepared by the Utah Education Policy Center December 2020

THE UNIVERSITY OF UTAH
UTAH EDUCATION POLICY CENTER

THE UNIVERSITY OF UTAH
UTAH EDUCATION POLICY CENTER

Bridging Research, Policy, and Practice

The Utah Education Policy Center (UEPC) is an independent, non-partisan, not-for-profit research-based center at the University of Utah founded in the Department of Educational Leadership and Policy in 1990 and administered through the College of Education since 2007. The UEPC mission is to bridge research, policy, and practice in public schools and higher education to increase educational equity, excellence, access, and opportunities for all children and adults.

The UEPC informs, influences, and improves the quality of educational policies, practices, and leadership through research, evaluation, and technical assistance. Through our research, evaluation, and technical assistance, we are committed to supporting the understanding of whether educational policies, programs, and practices are being implemented as intended, whether they are effective and impactful, and how they may be improved and scaled-up, and become sustainable.

Please visit our website for more information about the UEPC.

http://uepc.utah.edu

Andrea K. Rorrer, Ph.D., Director
Phone: 801-581-4207
andrea.rorrer@utah.edu

Cori Groth, Ph.D., Associate Director
Phone: 801-581-4207
cori.groth@utah.edu

Follow us on Twitter: @UtahUEPC

Citation: Su, Y., Rorrer, A., Owens, R., Pecsok, M., Moore, B., & Ni, Y. (2020). K-12 Mathematics Personalized Learning Program 2018-19 Student Outcomes, Evaluation Addendum for Year 3. Salt Lake City, UT: Utah Education Policy Center.

Table of Contents

List of Tables

List of Figures

Introduction

The Utah STEM Action Center (STEM AC) was created in 2013 by the Utah legislature (House Bill 139[1]) and expanded by House Bill 150[2] STEM AC, which is now a division of the Utah Department of Heritage & Arts, advances STEM Education in Utah by providing opportunities for students, educators, and educational learning partners to expand STEM learning (e.g., digital learning tools, professional development, STEM programs). One program offered by the STEM AC is the K-12 Math Personalized Learning Software Grant program, which is the focus of this evaluation amendment.

K-12 Math Personalized Learning Software Grant Overview

The K-12 Math Personalized Learning Software Grant program[3] provides access to approved math personalized learning software vendor licenses to teachers and students in Utah schools and districts, Schools apply for grant funds through the STEM AC to purchase these licenses. The overall program has two intended outcomes for the grant program. One outcome is to increase student awareness, engagement, interest, and perceived utility of math. The other outcome of the grant is to improve math literacy, which can be measured by increases in student performance on standardized math test.

Math Personalized Learning Software Program Evaluation Overview

The Utah STEM Action Center math personalized learning software program has been evaluated by the Utah Education Policy Center (UEPC) since 2016. To date, the UEPC has evaluated outcomes for the STEM AC math personalized learning software programs across three domains:

- Teacher knowledge, practice, and outcomes
- Student attitudes and perceptions
- Student learning as measured by standardized math test scores

As an Addendum to the Year 3 Evaluation, this report addresses the last domain of these three outcomes—measuring student learning through math standardized test. For the most recent evaluations of the other two domains, we encourage the reader to explore our annual reports: *Teaching Mathematics with Technology in Utah: An Evaluation of Teacher Knowledge, Practices, and Outcomes Using Mathematics Personalized Learning Software (*Onuma et al., 2020*),* and *Learning Mathematics with Technology in Utah: An Evaluation of Student Attitudes and Perceptions of Math Personalized Learning Software (*Auletto et al., 2020).

Purpose of the Evaluation

During the 2018-2019 school year, Utah schools experienced issues with the state's end-of-year assessments—the Readiness Improvement Success Empowerment (RISE)—which caused a

[1] https://le.utah.gov/~2013/bills/static/HB0139.html
[2] https://le.utah.gov/~2014/bills/static/HB0150.html
[3] K-12 Math Personalized Learning Software Grant, https://stem.utah.gov/grants/k-12-math-personalized-learning-software-grant/

delay in the availability of student outcome data, and subsequently delayed the analysis of the relationship between student use of math personalized learning software programs and student outcomes on the RISE math test. In the interim, the UEPC provided a supplemental longitudinal study of student outcomes impacted by math personalized learning software programs. The results of the longitudinal evaluation study, which uses a different methodology than this study, can be found in *Impact of K-12 Math Personalized Learning Software on Student Achievement* (Owens et al., 2020).

This Year 3 Addendum continues the pattern of evaluating the impact of student software use on 2018-2019 math achievement outcomes on the state's end-of-year math assessment (RISE). See previous evaluation report—*Utah STEM Action Center Program Evaluation: Academic Year 2018-19* (Eddings et al., 2019)—for immediate past results. Importantly, using propensity score matching, the analysis presented here reports results of a comparisons in math performance among students who used licenses provided by STEM AC and students who were at schools that did not receive software licenses through STEM AC. (See *Methodology* section for additional information). Moreover, the analysis presented here controls for student characteristics so that we can more precisely estimate the impact of the math personalized software use as an intervention. While analyses for previous evaluations utilized SAGE scores, Utah's end-of-year assessment through the 2017-2018 school year, this analysis examines RISE mathematics test outcomes. Although results are reported by vendor, this evaluation is about the personalized learning program offered in its totality and not rendering judgement on any given vendor.

Literature Review

In this section, we focus our review of literature on the context for increased use of personalized mathematics software and the factors that influence successful implementation of personalized learning software in teaching and learning and the subsequent outcomes of software use. Although there is some debate on the effectiveness of digital learning software for improving student learning under different contexts, many empirical research studies demonstrate that appropriate implementation of digital learning software is key to improving student math learning outcomes (Koedinger, McLaughlin, & Heffernan, 2010; Pilli & Aksu, 2013; Roschelle et al., 2010).

Mathematics performance in the United States often lags behind that of many other countries (Higgins et. al, 2019). Indeed, the proportion of US high school graduates who performed at an advanced level on the Program for International Student Assessment (PISA) is much lower than most comparable developed nations, including Switzerland, Belgium, and The Netherlands (Hanushek, et al, 2010). One strategy to address underperformance in math has been to increase the utilization of technology to supplement mathematics instruction (Higgins et al., 2019; Li & Ma, 2010). Since the 1990s, technology availability and utility as an instructional tool has dramatically evolved and increased (De Witte, Haelermans, & Rogge, 2015; Franklin, Orians, & Rorrer, 2016; Morgan, 2008). Recent events, such as the shift to remote learning during the COVID-19 pandemic, have further highlighted the importance of technology for both teaching and learning (Schaffhauser, 2020). Research (Cheung & Slavin, 2013; Murphy, 2016; Young, Gorumek, & Hamilton, 2018) has demonstrated benefits of using technology in mathematics education, including engaging student in the learning process, improvement in teacher-student interactions, enhancing students' higher-order problem solving techniques, and increasing student math achievement.

Effective implementation of digital learning software can be challenging (Al-Adwan, & Smedley, 2012; Franklin et al., 2016; Sarker et al., 2019). Studies suggest there are multiple factors that influence integration of education technology with classroom instruction, including the ability to use technology, difficulties with the integration of the technology in teaching, absence of expertise and insufficient technical support (e.g., Bauer & Kenton, 2005; Ertmer, 2005; Karagiorgi & Charalambous, 2004; Tatar, 2013). Teachers' pedagogical knowledge for appropriately integrating technology into curricula has been noted as critical to student learning (Hew & Brush, 2007; Hughes, 2005; Lawless & Pellegrino, 2007; Sarker et al., 2019). For instance, recent research demonstrates that student achievement increased when teachers adjusted their technology use to better align with curriculum and instruction (Callaghan et al., 2018). To this end, theoretical foundations on digital learning tool use in education, such as the Substitution, Augmentation, Modification, and Redefinition (SAMR) model (Puentedura, 2013) and the Technological Pedagogical and Content Knowledge (TPACK) Framework (Mishra & Koehler, 2006), provide us with a complete picture of the key elements that constitute effective implementation of educational technology and a productive approach to how they coordinate together to improve student learning experiences and outcomes.

Interestingly, empirical studies reveal that some software use patterns significantly improve student learning outcomes. For instance, Sharp and Hamil (2018) found that students who attempted a greater number of sessions on a web-adaptive digital resource had higher scores on a state math assessment. Similarly, DeWitt and colleagues (2015) argued that completing more lessons in a computer-assisted instruction program was related to higher scores on school tests and a national test. Heinrich et al. (2019) found that students who spent more time in their online courses and completed more activities per day had better course performance. These findings are particularly relevant for this evaluation.

Evaluation Questions

To examine the impact of math personalized learning software program use on student math achievement outcomes during the 2018-19 school year, we explore the following evaluation questions:

1. What is the relationship between the amount of weekly program use and RISE math proficiency while controlling for student characteristics?
2. What is the relationship between the amount of weekly program use and Student Growth Percentile (SGP) score while controlling for student characteristics?

In addition to the overall findings, results for these evaluation questions are disaggregated by vendor who participated in the STEM AC grant program.

Methodology

The focus of the current evaluation is to understand relationships between math personalized learning software program use and student math achievement outcomes on RISE math proficiency and SGP. Here we provided an overview of our data sources and data analysis procedures.

In order to gain a more precise understanding of the relationship between software use and math performance, we control for student characteristics in the regression analysis, including race/ethnicity, gender, qualification for Free and Reduced Lunch, English learner status, rural status, special education status, chronic absence status, mobile status, grade level, and previous year proficiency. Doing so allows us to more precisely estimate the impact of the intervention. The framework for our logistic and linear regression data analyses is presented in Figure 1, which includes the key indicator of weekly program use, controlling variables of student characteristics, and the two math achievement outcomes we examined—RISE math proficiency and SGP. Table 1 provides definitions for key variables introduced in Figure 1.

Figure 1. Data Analysis Framework

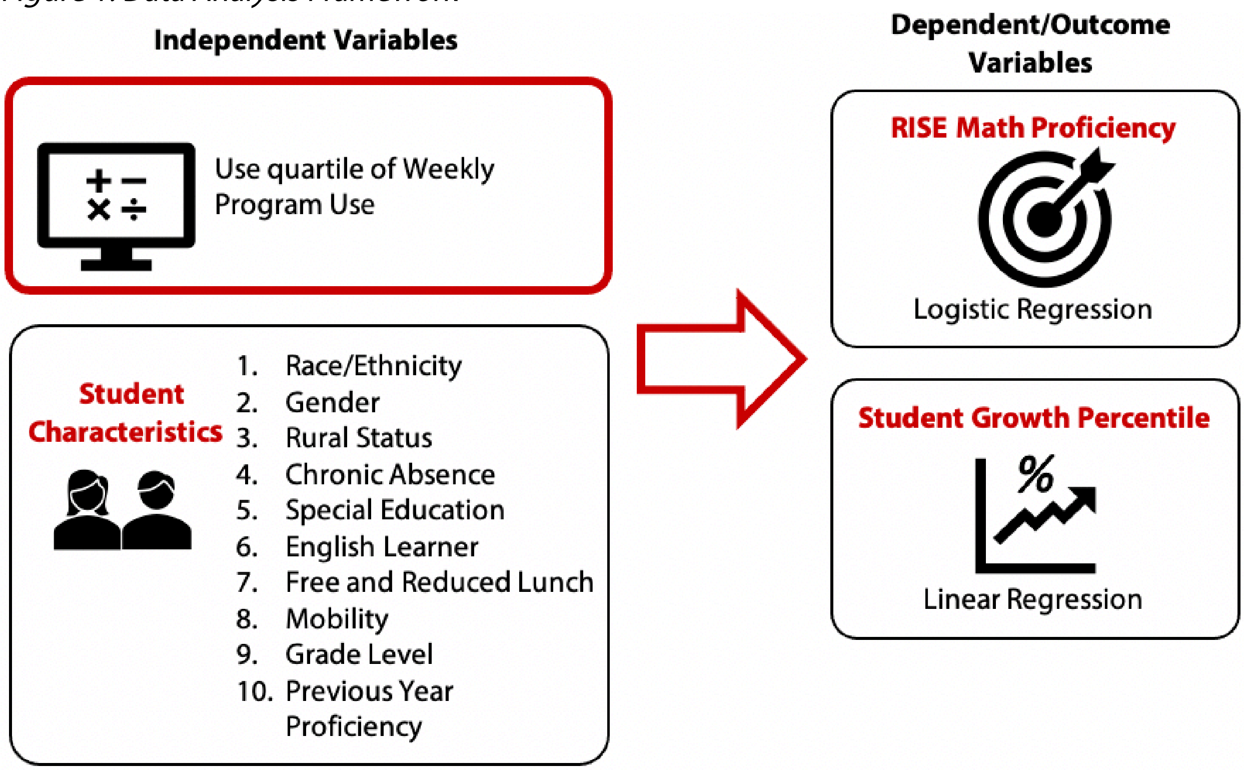


Table 1. Definition of Variables Used in Regression Models

Variable	Definition
Weekly Program Use	Represented as four quartiles by dividing the rank-ordered users' weekly time usage into four equal parts (each part takes 25% of user sample size). Students in the first quartile used software the fewest minutes per week and students in the fourth quartile used software the most minutes per week.
Race/Ethnicity	A categorical variable, including White, Black or African American, Asian, American Indian and Alaskan Native, Native Hawaiian and Pacific Islander, or multiple races.
Gender	A binary variable, including male and female.
Rural Status	A binary variable, including rural and non-rural. Students whose schools were designated as rural were categorized as rural. Otherwise, they were categorized as non-rural.
Chronic Absence	A binary variable. Students were considered to be chronically absent if they were absent for 10% or more of their total membership days in schools in which they had the highest number of attendance days. Otherwise, students were considered as non-chronically absent.
Special Education	A binary variable, including students who received special education services and those who did not.
English Learner	A binary variable, including English learners and native English speakers
Free and Reduced Lunch	A binary variable, including eligible and non-eligible status.
Mobility	A binary variable. Students were classified as mobile if they attended more than one school in a given school year. Otherwise, they were classified as non-mobile.
RISE Math Proficiency	A binary variable based on end-of-year statewide summative math test scores, indicating whether a student is on track towards college and career readiness in math by categorizing scaled scores into two levels, proficient or not proficient.
Student Growth Percentile (SGP) Score	A measure of student growth or improvement from the previous year summative math score relative to peers. The range of a SGP score is from 1 (lowest growth) to 99 (highest growth). For example, a SGP score of 52 means the student showed more growth or greater improvement than 52% of other students who performed similarly to them the previous year.

Data Sources

The data used for this analysis were from two sources. Student program use data were provided by the five participating software vendors (ALEKS, Imagine Math, i-Ready, Mathspace, ST Math) and two pilot programs (Dreambox and MobyMax)[4]. Each vendor provided monthly software usage reports to the UEPC between September 2018 and May 2019. Student data, including student characteristics and achievement data, were made available via a Data Share Agreement between the UEPC and the Utah State Board of Education (USBE)[5]. See Appendix B for further details about how the UEPC handles data.

[4] Data from IXL Learning, an eighth vendor participating in the STEM AC grant program, were not included because the sample size of records did not meet the threshold for the analysis.

[5] The views expressed in this report are those of the authors and are not necessarily the USBE's nor endorsed by the USBE.

Consistent with our methodology for the longitudinal report (Owens, et al., 2020), the primary sample for this evaluation consists of Utah public school students who used one of the math personalized learning software programs funded or piloted through STEM AC during 2018-19 school year. The RISE math test was offered from grade 3 and above. A dataset for analysis was created using the following criteria: vendor login and school year information, name, school identification and school year matching algorithm. The following data exclusion rules were applied to establish the sample and dataset:

- Students in kindergarten through second grade.
- Vendor data that we were unable to match to USBE records.
- Records that had excessively high time usage data (e.g. about 15 hours/week).
- Software user records with less than one minute of reported program use or zero logins. They were considered as neither software users nor students from comparison schools.

Math personalized learning software program users and students from comparison schools were compared in this study. The software users (n=104,437) were defined as the students who logged into one of the seven math software programs at least one minute during the school year (September 2018 - May 2019). Students from comparison schools (n=218,665) were defined as the students who came from propensity score matched schools (See *Data Analysis* for details of matching procedures) and did not generate any log data reported by any of the seven vendors whose licenses were funded (ALEKS, Imagine Math, i-Ready, Mathspace, and ST Math) or piloted (DreamBox and MobyMax) by STEM AC during the 2018-19 school year[6]. Table 2 shows the sample of math personalized learning software users and their matched students from comparison schools by vendor included in the analysis.

Table 2. Sample Size by Vendor

	ALEKS	DreamBox	Imagine Math	i-Ready	Mathspace	MobyMax	ST Math	Total
Software Users	58,858	475	14,510	16,294	2,105	431	11,764	104,437
Comparison School Students	89,741	771	28,915	32,989	11,319	13,042	41,888	218,665

Data Analysis

Here we provide an overview of data analysis process used for this evaluation. First, we describe the reasons why we use propensity score matching and how we apply this technique in our analysis. Then, we explain the context of using regression analysis and the steps that we conduct in the current study.

[6] There is no data related to whether or not students from comparison schools used any of these seven software vendors from sources other than the STEM AC.

Propensity Score Matching

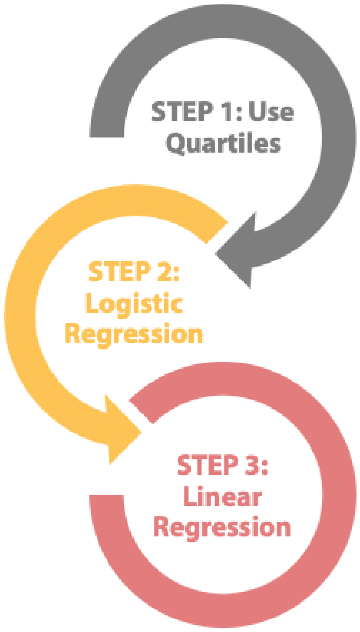
Data analysis for this evaluation included propensity score matching methods. Propensity score matching technique is used to minimize the chances that differences in testing outcomes between students who used the math personalized learning software program through a STEM AC license and students from comparison schools who did not used it through a STEM AC license were the result of factors, such as student characteristics, other than the math personalized learning software program use. With this technique, math performance of students from comparable schools serves as the baseline to which program users are compared. Note that we did not use math performance of all Utah students who did not use the math personalized learning software programs through STEM AC licenses as the baseline. A benefit of this approach is to better assess the relationships between program use through a STEM AC license and RISE math proficiency and SGP by reducing the selection bias in estimating the effects of software programs (Adelson, 2013).

Propensity score matching in this study requires two steps. First, schools associated with the seven software vendors whose licenses were funded or piloted by STEM AC were identified and proportions of student characteristics at the school level were calculated, including race/ethnicity, gender, rural status, chronic absence status, special education status, English learner status, qualification for Free and Reduced Lunch, and mobility status. Second, based on the calculated proportions, comparable schools which were not supported through the STEM AC grant program were identified.

Regression Analysis

Logistic and linear regression models were used to examine the relationships between program use through a STEM AC license and math achievement outcomes, including RISE math proficiency and SGP score. Use of regression models in the analysis allows us to control for student characteristics at the student level when estimating the relationships between program use through a STEM AC license and student math achievement outcomes (Long & Mustillo, 2018). The primary reason to control for student characteristics is to more precisely estimate the impact of the intervention (see Figure 1 for data analysis framework). The data analysis procedures are outlined in Figure 2.

Figure 2. Data Analysis Procedures



STEP 1: Weekly program use was divided into quartiles based on the average amount of time software users spent on using it each month of use. Due to the variation of minutes each software program was used, the cutoff values for the quartiles were unique for each vendor.

STEP 2: Due to the dichotomous nature of the test proficiency outcome, logistic regression was used to estimate the relationship between weekly program use and student RISE math test proficiency. In this model, our sample included software users and matched students from comparison schools. At the same time, student characteristics were included as control variables in the model. We also investigated how they correlated with the student RISE math proficiency as a secondary interest.

STEP 3: A linear regression model was used to explore the relationship between weekly program use and SGP score due to the continuous nature of the SGP score outcome. Our sample in this model was exactly the same as STEP 2 within an individual software vendor. Like the logistic regression model, we also accounted for the same set of student characteristics variables as control variables and examined their relationships with SGP score.

As Figure 2 shows, the approach to define quartiles of weekly software program use was different than previous evaluations because we can have an in-depth understanding of varied program time use patterns across software vendors. Instead of cutting quartiles with a universal standard for all software vendors, the quartiles in this study are defined in terms of the program time use for each software vendor to capture different time use patterns across software vendors. See *Addendum to the 2017-18 STEM Action Center Program Evaluation* (2019) for previous quartile definition.

Furthermore, both logistic and linear regression models used in Step 2 and Step 3 in this study treat the math performance of matched students from comparable schools as the baseline to compare to that of software users, which is different from the longitudinal study that used math performance of the lowest quartile of program use as the baseline to compare to each of other three quartiles. See *Impact of K-12 math personalized learning software on student achievement* (Owens et al., 2020) for comparison with longitudinal results.

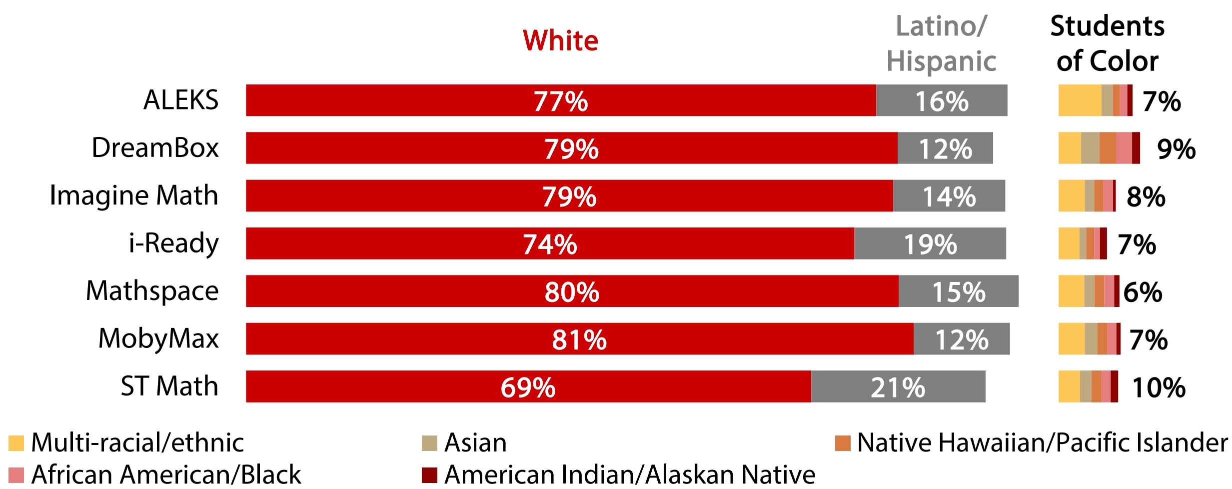
Results

In this section, we first describe the demographic characteristics of math personalized software users and students from comparison schools. To provide context, the information is disaggregated by vendor. Then, we present findings by evaluation question. As a secondary interest, we provide findings about how controlled student characteristic variables related to key outcomes. Similarly to student demographic characteristics, we have disaggregated results by software vendor. Again, note that while information is disaggregated by software vendor, we are not evaluating the efficacy of any given vendor's product. Instead, this evaluation is on the overall impact of the math personalized learning software programs funded or piloted by STEM AC.

Demographics of STEM AC Math Personalized Learning Software Users and Students in Comparison Schools

Figures 3-4 and Table 3-5 describe characteristics of students in schools who participated in the STEM AC math personalized learning software program and those matched from comparison schools who did not use them through a STEM AC license. We report percent of software users and students from comparison schools in one set of tables and figures because after propensity score matching, student characteristics were similar between software users and students from comparison schools for each vendor. Figure 3, Table 3, and Table 4 provide a summary of student characteristics. Table 5 and Figure 4 contain descriptive information about weekly software program use.

Figure 3. Race/Ethnicity of Program Users and Students from Comparison Schools Disaggregated by Vendor



As Figure 3 shows, a majority of students (69-81%) were White and 12-21% were Latino/Hispanic. Other Students of Color accounted for no more than 10% of program users or their matched students from comparison schools per vendor included in the analysis, including Multi-racial/ethnic, Asian, Native Hawaiian/Pacific Islander, African American/Black, and American Indian/Alaskan Native.

Table 3. Student Characteristics of Software Users and Students from Comparison Schools Disaggregated by Vendor

	ALEKS	DreamBox	Imagine Math	i-Ready	Mathspace	MobyMax	ST Math
Gender							
Male	51%	50%	51%	51%	51%	51%	51%
Female	49%	50%	49%	49%	49%	49%	49%
Other Characteristics							
Previous Year Proficient	43%	53%	49%	45%	44%	45%	45%
Free and Reduced Lunch	33%	32%	27%	36%	37%	21%	39%
English Learners	7%	5%	7%	10%	8%	4%	15%
Rural	12%	0%	9%	9%	17%	2%	10%
Special Education	12%	16%	14%	14%	14%	13%	14%
Chronic Absence	10%	11%	11%	12%	11%	7%	11%
Mobile	3%	3%	3%	4%	3%	3%	4%

Table 3 describes the demographic and background information of STEM AC math personalized learning software users and matched students from comparison schools included in the analysis, reported by vendor. Overall, the following observations can be made about the sample:

- Male and female students were near evenly represented.
- Approximately half of students were proficient on the previous year's SAGE math test.
- About one-third of students qualified for Free and Reduced Lunch.
- Only about one-tenth of students were English learners.
- Rural school students accounted for 17% or fewer.
- Students receiving special education services accounted for about 15% of the students in our analyses.
- Approximately one-tenth of students were chronically absent from classes during the 2018-19 school year.
- Only about 3% of students attended more than one school during the 2018-19 school year.

Table 4. Math Personalized Software Users and Students from Comparison Schools Disaggregated by Grade Level and Vendor

Vendor	3rd	4th	5th	6th	7th	8th	9th	10th	11th
ALEKS									
Software Users	-	5%	7%	12%	22%	21%	20%	12%	-
Students from Comparison School	<1%	14%	15%	15%	14%	14%	13%	15%	<1%
DreamBox									
Software Users	-	45%	50%	5%	-	-	-	-	-
Students from Comparison School	-	39%	42%	19%	-	-	-	-	-
Imagine Math									
Software Users	<1%	28%	29%	24%	9%	6%	4%	-	-
Students from Comparison School	<1%	25%	26%	25%	8%	5%	4%	6%	<1%
i-Ready									
Software Users	<1%	32%	32%	21%	8%	8%	<1%	-	-
Students from Comparison School	<1%	23%	23%	22%	11%	7%	7%	8%	<1%
Mathspace									
Software Users	-	17%	12%	32%	16%	8%	8%	7%	-
Students from Comparison School	<1%	18%	19%	21%	12%	11%	8%	11%	<1%
MobyMax[7]									
Software Users	-	46%	34%	17%	1%	1%	-	-	-
Students from Comparison School	<1%	15%	15%	18%	9%	9%	5%	29%	<1%
ST Math									
Software Users	-	39%	34%	21%	3%	3%	<1%	<1%	-
Students from Comparison School	<1%	23%	24%	21%	11%	10%	7%	3%	<1%

Table 4 shows the distribution of software users and students from comparison schools by grade level and software vendor. The software program availability varies from one another. The majority of students for most vendors were in grades 4 to 6.

[7] Propensity score matching did not match percentages of grade level between software users and students from comparison schools, but grade level was later controlled in regression analysis.

Table 5. Minutes of Weekly Program Use among Software Users

Vendor	Quartile 1	Quartile 2	Quartile 3	Quartile 4
ALEKS	<22	22-41	>41-70	>70
DreamBox	<18	18-30	>30-49	>49
Imagine Math	<14	14-23	>23-37	>37
i-Ready	<21	21-33	>33-46	>46
Mathspace	<6	6-12	>12-33	>33
MobyMax	<3	3-6	>6-13	>13
ST Math	<18	18-31	>31-50	>50

ALEKS users spent the most time

MobyMax users spent the least

Figure 4. Median Minutes of Weekly Program Use by Quartile for Software Users Disaggregated by Vendor

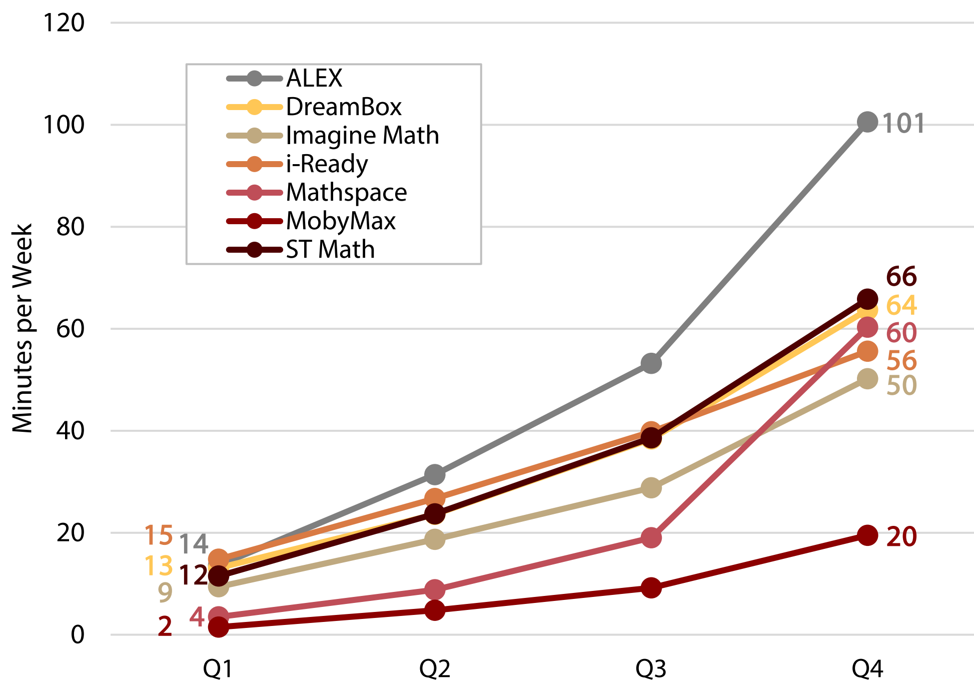


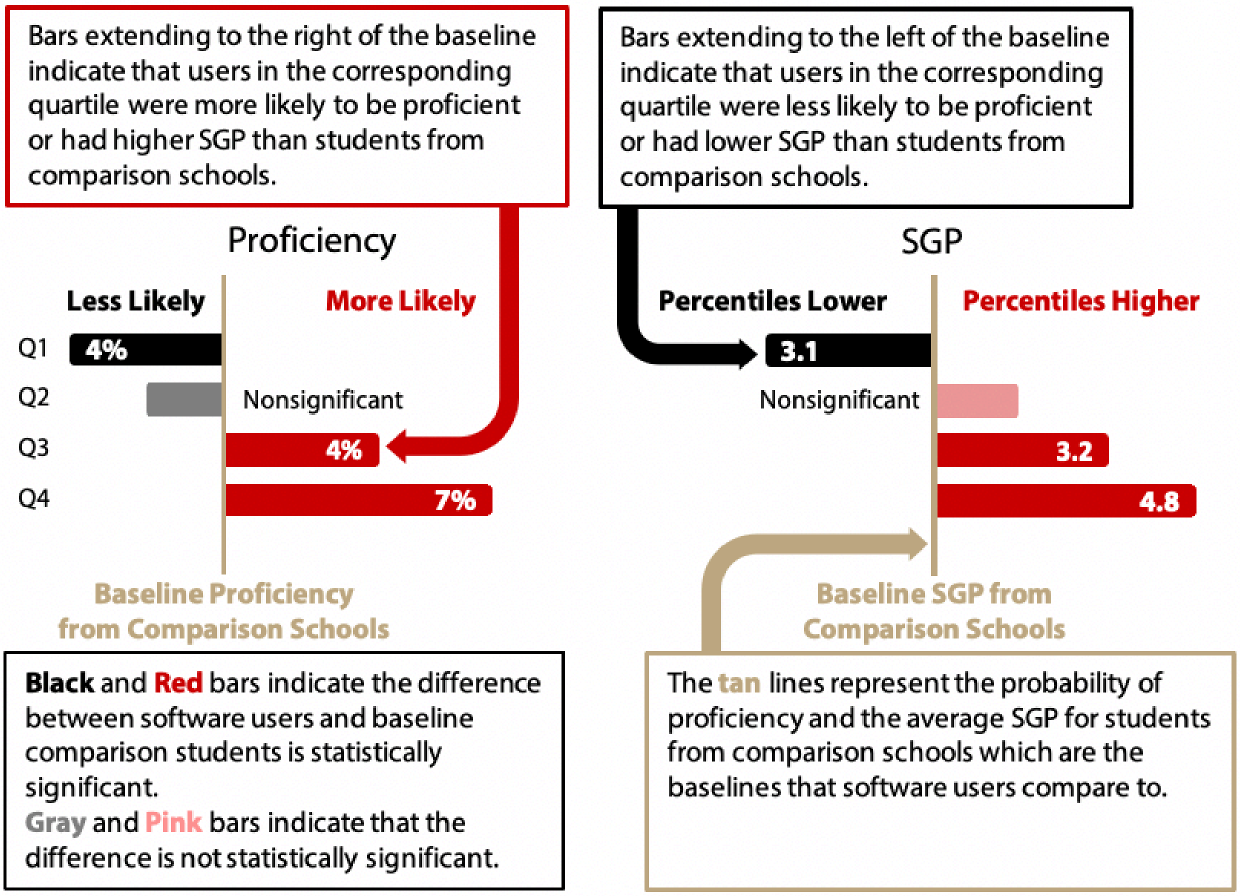
Table 5 and Figure 4 describe Quartiles of Math Software Program Use. Table 5 depicts the average weekly use by quartile across the math personalized learning software programs funded or piloted by STEM AC. The amount of time students spent weekly using software was a key indicator of RISE math proficiency and SGP investigated in this evaluation. Based on the weekly program use within each vendor, software users were grouped into four quartiles. Each quartile included a group of 25% of users rank-ordered by average minutes per week of program use (See Table 2 for a more detailed definition). Figure 4 provides the median of each quartile of weekly program use across all vendors. As indicated in both Table 5 and Figure 4, ALEKS users generally spent more minutes per week using software than software users of the other software programs. In contrast, students used MobyMax the least time per week.

Relationships Between Math Personalized Learning Software Use and RISE Math Proficiency and Student Growth Percentile

In this section, we report on the relationship between the amount of weekly program use and RISE proficiency and the relationship between the amount of weekly program use and Student Growth Percentile (SGP) score. We describe how to interpret regression results and then report findings according to evaluation questions by each vendor. Furthermore, the Appendix A provides detailed regression results for each vendor, which includes in-depth statistical results for all variables involved in the analysis in addition to the key findings we present below.

Figure 5 is provided to offer guidance for interpreting the results for each vendor. Results are presented to illustrate the difference in performance between STEM AC math personalized software program users and their peers from comparable schools who did not use them through a STEM AC license. This information is reported by vendor software use quartile (See Table 5. Minutes of Weekly Program Use among Software Users). Moreover, additional information is provided to illustrate whether the differences in RISE math proficiency and the SGP score were statistically significant. Figures 6-12 provide a summary of primary regression results for each vendor. In these figures, the proficiency bar chart on the left shows the estimates of differences in probability of RISE math proficiency between STEM AC program users and students from comparison schools who did not use software programs through a STEM AC license, while the SGP bar chart on the right illustrates the estimates of the differences in SGP scores between these two groups.

Figure 5. How to Read Figures with Regression Results for Each Vendor

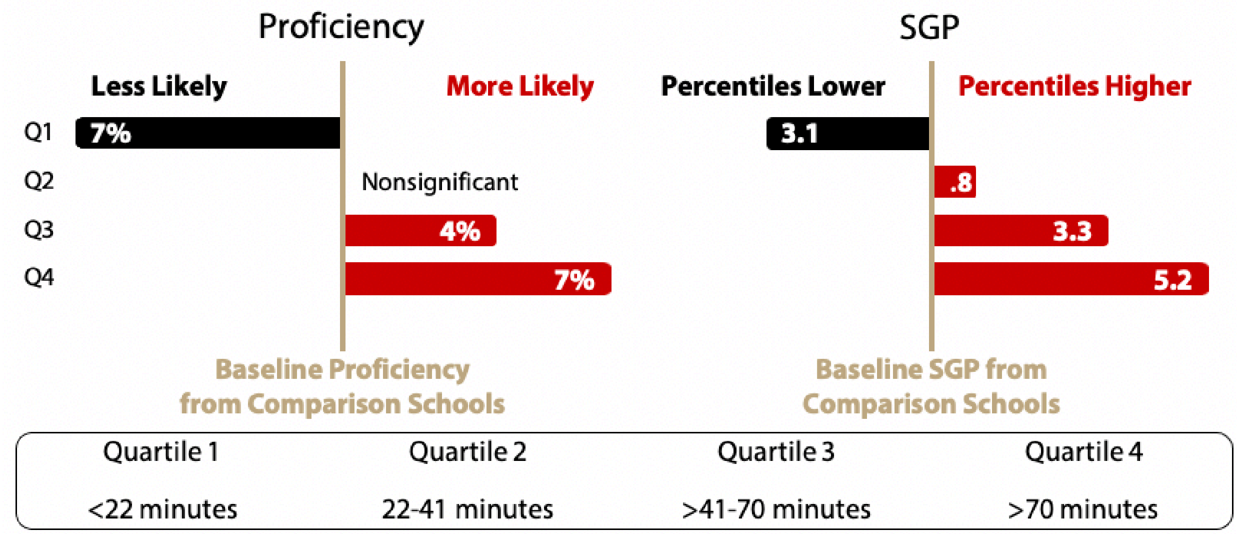


Statistical Significance
In statistics, significance refers to whether any differences observed between groups are "real" or whether they are simply due to chance.

In addition to the estimated differences in outcomes between STEM AC software users and students from comparison schools who did not use software programs through STEM AC license that are depicted in Figures 6-12, we also estimated regression models that consider how student characteristics related to RISE math proficiency and SGP (see tables in Appendix A). It is important to note, however, that those group differences are not related to vendor use. A baseline group for comparison was assigned for each student characteristic, as noted in the methodology section. For example, the baselines for gender and race were male and White, respectively. Other baselines used as student characteristics included non-proficient students on the previous year's math test, students who did not qualify for Free and Reduced Lunch, students who were not English learners, students from non-rural schools, students who did not receive special education services, students who were not chronically absent, students who were not mobile, and students who were in grade 4 (we used grade 4 because third-grade students did not have prior math test scores).

ALEKS Results

Figure 6. ALEKS Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly ALEKS Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there were positive relationships between weekly program use of ALEKS in Quartiles 3 and 4 (i.e., weekly time usage >41 minutes) and student RISE math proficiency. This finding was statistically significant ($p<.05$). In addition, there was a negative relationship between weekly program use of ALEKS in Quartile 1 (i.e., less than 22 minutes per week of use) and RISE math proficiency. This finding was statistically significant ($p<.05$). No statistically significant difference was found between program users of ALEKS in Quartile 2 (i.e., between 22-41 minutes per week of use) and comparison students as it relates to RISE math proficiency. See Appendix A, Table 7 for further details.

As illustrated in Figure 6, students who used the ALEKS program between 41 and 70 minutes per week (Quartile 3) had a 4% higher probability of RISE math proficiency than students from comparison schools who did not use it through a STEM AC license. The finding was statistically significant ($p<.05$). For students who used the ALEKS program in excess of 70 minutes per week (Quartile 4), their probability of being proficient on the RISE math assessment was 7% higher than that of students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Software users in Quartile 1 had a 7% lower probability of being RISE math proficient than students from comparison schools that was statistically significant ($p<.05$), while software users in Quartile 2 (i.e., 22-41 minutes per week of use) were not statistically significant different than student from comparison schools who did not use ALEKS through a STEM AC license for being proficient on the RISE math assessment. In summary, in the case of ALEKS, only students with weekly program use greater than 41 minutes (i.e., Quartiles 3 and 4) had a statistically significant and better chance of being proficient on the RISE math assessment than students from comparison schools who did not use the ALEKS program through a STEM AC license. In addition, other statistically significant and positive relationships with RISE math proficiency were found ($p<.05$) for:

- Asian students who had a 7% higher probability of being proficient on the RISE math assessment than White students.
- Students who were proficient on the previous year's SAGE test had a 36% higher probability of being proficient on the RISE math assessment in the 2018-19 school year than students who were not proficient on previous year's SAGE test.
- Rural school students who had a 3% higher probability of being proficient on the RISE math assessment than students in non-rural schools.

See Table 7 in Appendix A for a full list of relationships between student characteristics and RISE math proficiency for ALEKS.

Relationships Between Weekly ALEKS Program Use and SGP Score. Using a linear regression model, our analysis indicates that there were positive relationships between weekly program use in Quartiles 2 to 4 (i.e., weekly time usage >=22 minutes) and SGP score. This finding was statistically significant ($p<.05$). See Appendix A, Table 8 for further details. As Figure 6 shows, students who used the ALEKS program through a STEM AC license at least 22 minutes per week (Quartiles 2-4), had higher SGP scores respectively, .8, 3.3, and 5.2, than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). For students who used the ALEKS program through a STEM AC license less than 22 minutes per week (Quartile 1), they had 3.1 lower SGP scores than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). In the case of ALEKS, students with weekly program use at least 22 minutes (Quartiles 2-4) had statistically significant and higher SGP scores than students from comparison schools who did not use the ALEKS program through a STEM AC license.

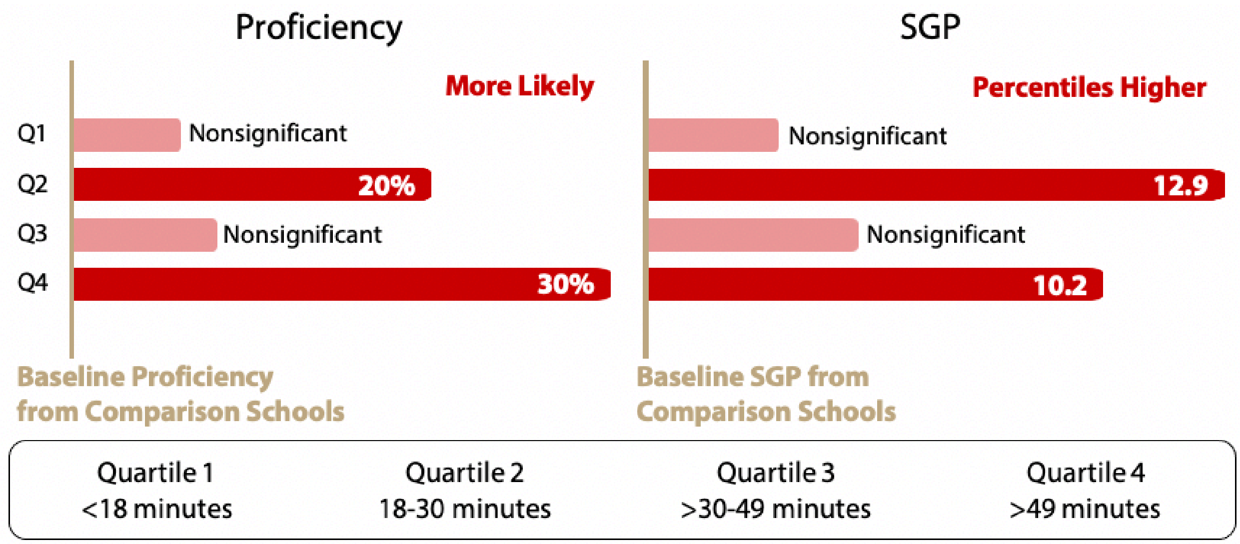
In addition, other statistically significant and positive relationships with SGP scores were found ($p<.05$) for:

- Asian students who had 3.5 higher SGP scores than White students.
- Rural school students who had SGP scores that were 1.4 higher than those of students in non-rural schools.

In Appendix A, Table 8 for a full list of relationships between student characteristics and SGP scores for ALEKS.

DreamBox Results

Figure 7. DreamBox Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly DreamBox Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there were positive relationships between program use of DreamBox in Quartile 2 (i.e., weekly time usage 18-30 minutes) and 4 (i.e., weekly time usage >49 minutes) and student RISE math proficiency. This finding was statistically significant ($p<.05$). See Appendix A, Table 9 for further details. As illustrated in Figure 7, students who used the DreamBox program between 18 and 30 minutes per week through a STEM AC license (Quartile 2) had a 20% higher probability of being proficient on the RISE math assessment than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). For students who used the DreamBox program through a STEM AC license in excess of 49 minutes per week (Quartile 4), their probability of being proficient on the RISE math assessment was 30% higher than that of, as compared to students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). In contrast, students who used the DreamBox program through a STEM AC license in Quartile 1 (i.e., less than 18 minutes per week of use) or 3 (i.e., between 30 and 49 minutes per week of use) also had respectively 6% and 8% higher probability of being proficient on the RISE math assessment as compared to students from comparison schools who did not use it through a STEM AC license. This finding, however, was not statistically significant. In summary, in the case of Dreambox, only students with weekly program use between 18 and 30 minutes or greater than 49 minutes had a statistically significant and better chance of being proficient on the RISE math assessment.

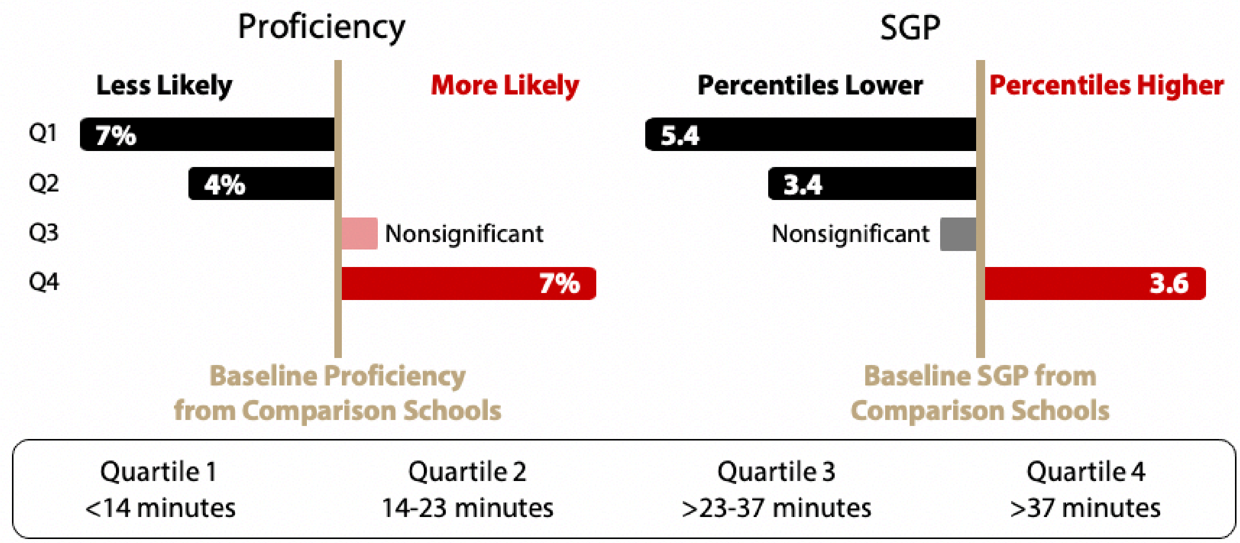
In addition, another statistically significant and positive relationship with RISE math proficiency was found for students who were proficient on the previous year's SAGE test. Results suggest that these students had a 33% higher probability of being proficient on the RISE math proficient assessment in the 2018-19 school year than students who were not proficient on the previous year's SAGE test. See Table 9 in Appendix A for a full list of relationships between student characteristics and RISE math proficiency for DreamBox.

Relationships Between Weekly DreamBox Program Use and SGP Score. Using a linear regression model, our analysis indicates that there were positive relationships between program use of DreamBox in Quartile 2 (i.e., weekly time usage 18-30 minutes) and 4 (i.e., weekly time usage >49 minutes) and SGP score. This finding was statistically significant ($p<.05$). See Appendix A, Table 10 for further details. As Figure 7 shows, for students who used the DreamBox program through a STEM AC license between 18 and 30 minutes (Quartile 2) or greater than 49 minutes per week (Quartile 4), they had higher SGP scores, respectively 12.9 and 10.2, than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the DreamBox program through a STEM AC license in Quartile 1 (i.e., less than 18 minutes per week of use) or 3 (i.e., between 30 and 49 minutes per week of use) also had higher SGP scores compared to students from comparison schools who did not use it through a STEM AC license. This finding, however, was not statistically significant. In summary, in the case of Dreambox, students with weekly program use between 18 and 30 minutes or greater than 49 minutes had statistically significantly higher SGP scores than students from comparison schools who did not use the Dreambox program through a STEM AC license.

In addition, another statistically significant and positive relationship with SGP scores was found for 6th graders who had 14.7 higher SGP scores than 4th graders ($p<.05$). See Table 10 in Appendix A for a full list of relationships between student characteristics and SGP score for DreamBox.

Imagine Math Results

Figure 8. Imagine Math Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly Imagine Math Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there was a positive relationship between program use in Quartile 4 (i.e., weekly time usage >37 minutes) and student RISE math proficiency. This finding was statistically significant ($p<.05$). See Appendix A, Table 11 for further details. As illustrated in Figure 8, students who used the Imagine Math program through a STEM AC license in excess of 37 minutes per week (Quartile 4) had a 7% higher probability of being proficient on the RISE math assessment than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the Imagine Math through a STEM AC license 23 minutes or less per week (Quartiles 1 and 2) had respectively 7% and 4% lower probability of being proficient on the RISE math assessment than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the Imagine Math program through a STEM AC license between 23 and 37 minutes per week (Quartile 3) had a 1% higher probability of being proficient on the RISE math assessment than students from comparison schools who did not use it through a STEM AC license. This finding was not statistically significant ($p<.5$). In summary, in the case of Imagine Math, only students with weekly program use greater than 37 minutes (Quartile 4) had a statistically significant and better chance of being RISE math proficient than students from comparison schools who did not use the Imagine Math program through a STEM AC license.

In addition, other statistically significant and positive relationships with RISE math proficiency were found ($p<.05$) for:

- Students who were proficient on the previous year's SAGE test. Results suggest that these students had a 34% higher probability of being proficient on the RISE math test during the 2018-19 school year than students who were not proficient on the previous year's SAGE test.
- Male students who had a 3% higher probability of being proficient on the RISE math assessment than female students.
- Rural school students who had a 4% higher probability of being proficient on the RISE math assessment than non-rural school students.
- Students in the 8th grade who had an 8% higher probability of being proficient on the RISE math assessment than 4th graders.

See Appendix A, Table 11 for a full list of relationships between student characteristics and RISE math proficiency for Imagine Math.

Relationships Between Weekly Imagine Math Program Use and SGP Score. Using a linear regression model, our analysis indicates that there was a positive relationship between program use in Quartile 4 (i.e., greater than 37 minutes per week of use) and SGP scores. This finding was statistically significant ($p<.05$). See Appendix A, Table 12 for further details. As illustrated in Figure 8, students who used Imagine Math through a STEM AC license in excess of 37 minutes per week (Quartile 4) had 3.6 higher SGP scores as compared to students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used Imagine Math through a STEM AC license 23 minutes or less per week (Quartiles 1 and 2) had respectively 5.4 and 3.4 lower SGP scores than students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the Imagine Math program through a STEM AC license between 23 and 37 minutes per week (Quartile 3) had 0.6 lower SGP scores than students who did not used the Imagine Math program through a STEM AC license. This finding, however, was not statistically significant. In summary, in the case of Imagine Math, only students with weekly program use greater than 37 minutes had statistically significant and higher SGP scores than students from comparison schools who did not use the Imagine Math program through a STEM AC license.

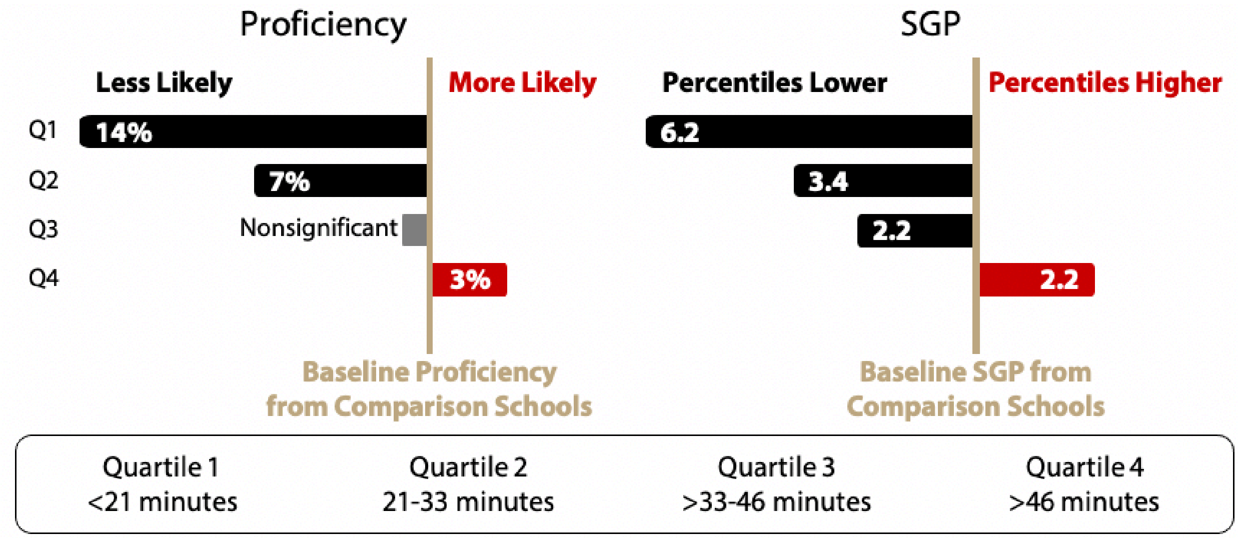
In addition, other variables tested had statistically significant and positive relationships with SGP scores ($p<.05$), including:

- Rural school students who had 2.5 higher SGP scores than students in non-rural schools.
- English learners who had 2 higher SGP scores than native English speakers
- 6th to 8th graders who had respectively 2.2, 2.2, and 5.3 higher SGP scores than 4th graders.

See Appendix A, Table 12 for a full list of relationships between student characteristics and SGP scores for Imagine Math.

i-Ready Results

Figure 9. i-Ready Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly i-Ready Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there was a positive relationship between program use in Quartile 4 (i.e., weekly time usage >46 minutes) and student RISE math proficiency. This finding was statistically significant ($p<.05$). See Appendix A, Table 13 for further details. As illustrated in Figure 9, students who used the i-Ready program through a STEM AC license in excess of 46 minutes per week (Quartile 4) had a 3% higher probability of being proficient on the RISE math assessment than students who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the i-Ready program 33 minutes or less per week (Quartile 1 and 2) had respectively 14% or 7% lower probability of being proficient on the RISE math assessment that were statistically significant ($p<.05$), while i-Ready user in Quartile 3 (i.e., between 33 and 46 minutes per week of use) had a 1% lower probability of being proficient on the RISE math assessment and that was not statistical significant ($p<.05$), as compared to students from comparison schools who did not use it through a STEM AC license. In summary, in the case of i-Ready, only students with weekly program use greater than 46 minutes (Quartile 4) had a statistically significant and higher probability of being proficient on the RISE math assessment than students from comparison schools who did not use the i-Ready program through a STEM AC license.

In addition, other statistically significant and positive relationships with RISE math proficiency were found ($p<.05$) for:

- Asian students who had a 7% higher probability of being proficient on the RISE math assessment than White students.

- Students who were proficient on the previous year's SAGE test. Results suggest that these students had a 34% higher probability of being proficient on the RISE math assessment during the 2018-19 school year than students who were not proficient on the previous year's SAGE test.
- Male students who had a 2% higher probability of being proficient on the RISE math assessment than female students

See Appendix A, Table 13 for a full list of relationships between student characteristics and RISE math proficiency for i-Ready.

Relationships Between Weekly i-Ready Program Use and SGP Score. Using a linear regression model, our analysis indicates that there was a positive relationship between program use in Quartile 4 (i.e., greater than 46 minutes per week of use) and SGP scores. This finding was statistically significant ($p<.05$). See Appendix A, Table 14 for further details. As illustrated in Figure 9, students who used the i-Ready program through a STEM AC license in excess of 46 minutes per week (Quartile 4) had 2.2 higher SGP scores than students who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the i-Ready program 46 minutes or less per week (Quartiles 1-3) had respectively 6.2, 3.4, & 2.2 lower SGP scores, as compared to students from comparison schools who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). In summary, in the case of i-Ready, only students with weekly program use greater than 46 minutes (Quartile 4) had a statistically significant and higher SGP scores than students from comparison schools who did not use the i-Ready program through a STEM AC license.

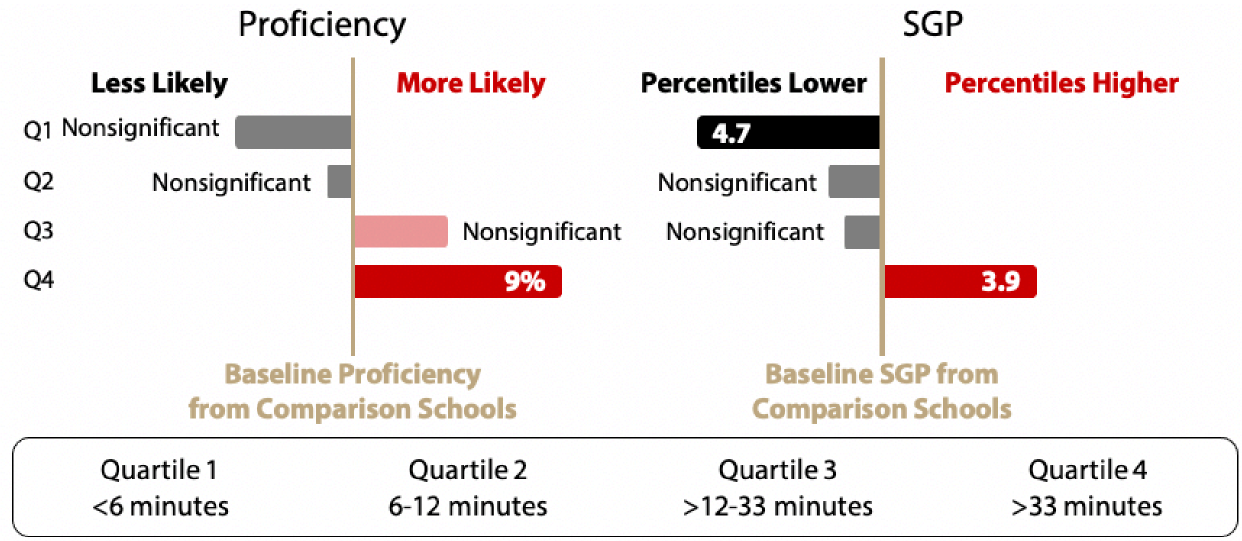
In addition, other statistically significant and positive relationships with SGP scores were found ($p<.05$) for:

- Asian students who had 4 higher SGP scores than White students.
- English learners who had 1.3 higher SGP scores than native English speakers.

See Appendix A, Table 14 for a full list of relationships between student characteristics and SGP score for i-Ready.

Mathspace Results

Figure 10. Mathspace Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly Mathspace Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there was a positive relationship between program use in Quartile 4 (i.e., weekly time usage >33 minutes) and student RISE math proficient. This finding was statistically significant (p<.05). See Appendix A, Table 15 for further details. As illustrated in Figure 10, students who used the Mathspace program through a STEM AC license in excess of 33 minutes per week had a 9% higher probability of being proficient on the RISE math assessment than students who did not use it through a STEM AC license. This finding was statistically significant (p<.05). Students who used the Mathspace program 12 minutes or less per week (Quartile 1 and 2) had respectively 5% and 1 % lower probability of being proficient on the RISE math assessment, while software users in Quartile 3 (i.e., between 12 and 33 minutes per week of use) had a 4% higher probability as compared to students from comparison schools who did not use it through a STEM AC license. Note that the relationships between Quartiles 1-3 and RISE math proficiency were not statistically significant (p<.05). In summary, in the case of Mathspace, only students with weekly program use greater than 33 minutes (Quartile 4) had a statistically significant and higher probability of being RISE math proficient than students from comparison schools who did not use the Mathspace program through a STEM AC license.

In addition, other statistically significant positive relationship with RISE math proficiency was found (p<.05) for:

- Students who were proficient on the previous year's SAGE test. Results suggested that these students had a 35% higher probability of being proficient on the RISE math

assessment during the 2018-19 school year than students who were not proficient on previous year SAGE test.
- Rural school students who had a 8% higher probability of being proficient on the RISE math assessment than non-rural school students.

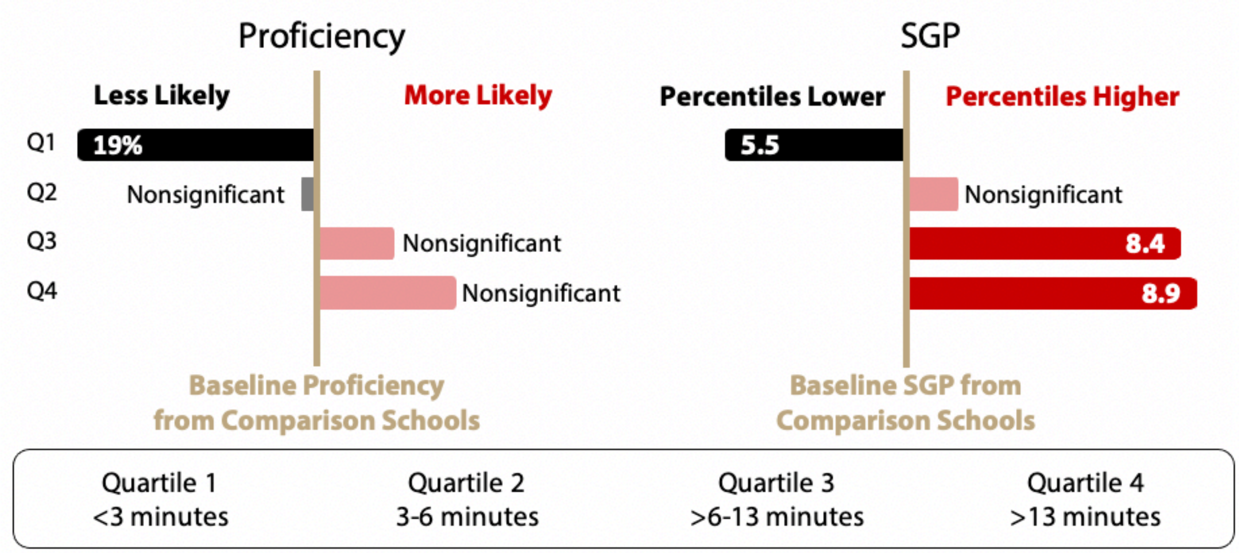
See Table 15 in Appendix A for a full list of relationships between student characteristics and RISE math proficiency for Mathspace.

Relationships Between Weekly Mathspace Program Use and SGP Score. Using a linear regression model, our analysis indicates that there was a positive relationship between program use in Quartile 4 (i.e., weekly time usage >33 minutes) and SGP scores. This finding was statistically significant ($p<.05$). See Appendix A, Table 16 for further details. As illustrated in Figure 10, students who used the Mathspace program through a STEM AC license in excess of 33 minutes per week (Quartile 4) had 3.9 higher SGP scores than students who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the Mathspace program through a STEM AC license less than 6 minutes per week (Quartile 1) had 4.7 lower SGP scores that was statistically significant ($p<.05$), while software users in Quartile 2 and 3 (i.e., between 6-33 minutes per week of use) had respectively 1.3 and 1 lower SGP scores that were not statistically significant ($p<.05$), as compared to students from comparison schools who did not use it through a STEM AC license. In summary, in the case of Mathspace, only students with weekly program use in excess of 33 minutes (Quartile 4) had statistically significant and higher SGP scores than students from comparison schools who did not use the Mathspace program through a STEM AC license.

In addition, we found that a statistically significant and positive relationship with SGP scores was for rural school students who had 5.5 higher SGP scores than non-rural school students ($p<.05$). See Table 16 in Appendix A for a full list of relationships between student characteristics and SGP score for Mathspace.

MobyMax Results

Figure 11. MobyMax Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly MobyMax Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there was no positive relationship that was statistically significant between program use in any quartile and student RISE math test proficiency ($p<.05$). See Appendix A, Table 17 for further details. As illustrated in Figure 11, though students who used the MobyMax program through a STEM AC license greater than 6 minutes per week (Quartile 3 and 4) had respectively 6% and 11% higher probability of being proficient on the RISE math assessment as compared to students from comparison schools who did not use it through a STEM AC license, however, they were not statistically significant ($p<.05$). Students who used the MobyMax program less than 3 minutes per week (Quartile 1) had a 19% lower probability of being proficient on the RISE math assessment which was statistically significant ($p<.05$), while software users in Quartile 2 (i.e., 3-6 minutes of use per week) had a 1% lower, however, not statistically significant ($p<.05$) probability of being proficient on the RISE math assessment as compared to students from comparison schools who did not use it through a STEM AC license. In summary, in the case of MobyMax, no matter how much time per week spent on using it, program users in any quartile did not have a higher probability of being proficient on the RISE math assessment which was statistically significant as compared to the students from comparison schools who did not use the MobyMax program through a STEM AC license.

In addition, other statistically significant and positive relationships with RISE math proficiency were found ($p<.05$) for:

- Students who were proficient on the previous year's SAGE test. Results suggested that these students had a 36% higher probability of being proficient on the RISE math test during the 2018-19 school year than students who were not proficient on the previous year's SAGE test.
- Male students who had a 3% higher probability of being proficient on the RISE math assessment than female students.
- 7th graders who had a 10% higher probability of being proficient on the RISE math assessment than 4th graders.

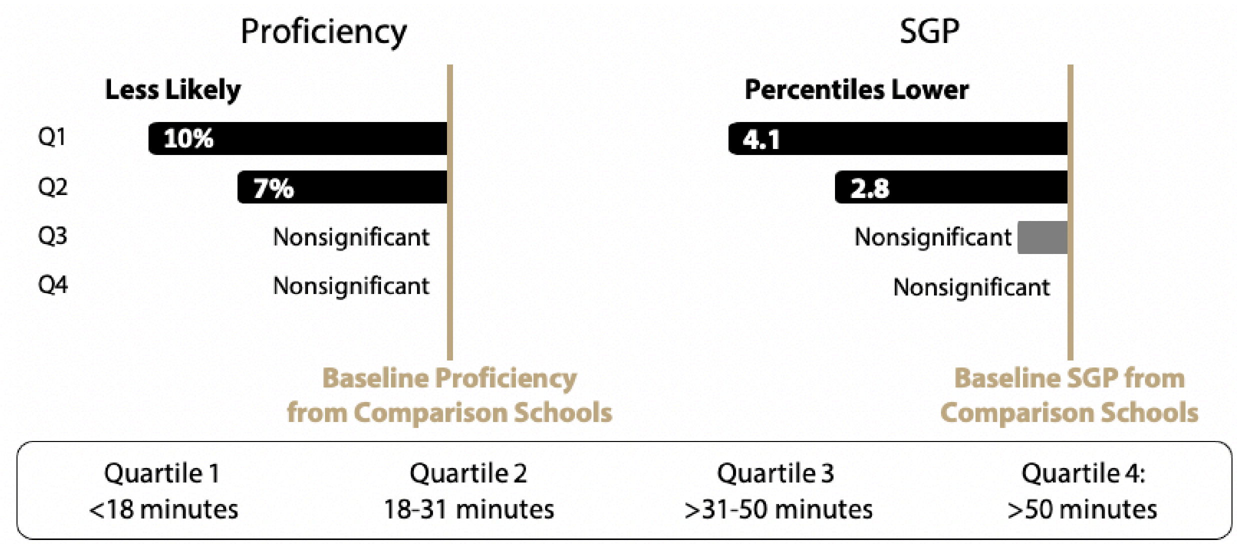
See Table 17 in Appendix A for a full list of relationships between student characteristics and RISE math proficiency for MobyMax.

Relationships Between Weekly MobyMax Program Use and SGP Score. Using a linear regression model, our analysis shows that there were positive relationships between program use in Quartiles 3 and 4 (i.e., weekly time usage >6 mins) and SGP scores. This finding was statistically significant ($p<.05$). See Appendix A, Table 18 for further details. As illustrated in Figure 11, students who used the MobyMax program greater than 6 minutes per week (Quartile 3 & 4) had respectively 8.4 and 8.9 higher SGP scores than students who did not use it through a STEM AC license. This finding was statistically significant ($p<.05$). Students who used the MobyMax program less than 3 minutes (Quartile 1) had 5.5 lower SGP scores which was statistically significant ($p<.05$), while program users in Quartile 2 (i.e., between 3 and 6 minutes per week of use) had 1.5 slightly higher SGP scores, however, not statistically significant, as compared to students who did not use it through a STEM AC license. In summary, in the case of MobyMax, only students with weekly program use greater than 6 minutes (Quartile 3 and 4) had statistically significant and higher SGP scores than students from comparison schools who did not use MobyMax through a STEM AC license.

In addition, another statistically significant positive relationship with SGP scores was found for 6th-9th graders who had respectively 3, 3.1, 5.1, and 6.1 higher SGP scores than 4th graders ($p<.05$). See Table 18 in Appendix A for a full list of relationships between student characteristics and SGP score for MobyMax.

ST Math Results

Figure 12. ST Math Results: The Relationships Between Weekly Program Use by Quartile and RISE Math Proficiency & SGP



Relationships Between Weekly ST Math Program Use and RISE Math Proficiency. Using a logistic regression model, our analysis demonstrates that there was no positive relationship that was statistically significant between program use in any quartile and student RISE math proficiency ($p<.05$). See Appendix A, Table 19 for further details. As illustrated in Figure 12, students who used the ST Math program 31 minutes or less per week (Quartile 1 and 2) had respectively 10% and 7% lower probability of being proficient on the RISE math assessment that were statistically significant ($p<.05$), while program users in excess of 31 minutes per week (Quartile 3 and 4) had no statistically significant different probability of being proficient on the RISE math assessment, as compared to students from comparison schools who did not use it through a STEM AC license ($p<.05$). In summary, in the case of ST Math, regardless of how much time per week spent on the ST math, program users in any quartile did not have statistically significant and higher probability of being proficient on the RISE math assessment as compared to the students from comparison schools who did not use the ST Math program through a STEM AC license.

In addition, other statistically significant and positive relationships with RISE math proficiency were found ($p<.05$) for:

- Asian students who had 8% higher probability of being proficient on the RISE math assessment than White students.
- Students who were proficient on the previous year's SAGE test. Results suggested that these students had a 34% higher probability of being proficient on the RISE math assessment on the RISE test during the 2018-19 school year than students who were not proficient on previous year SAGE test.

- Male students who had a 2% higher probability of being proficient on the RISE math assessment than female students.
- Rural school students who had 3% higher probability of being proficient on the RISE math assessment than students non-rural school students.

See Table 19 in Appendix A for a full list of relationships between student characteristics and RISE math proficiency for ST Math.

Relationships Between Weekly ST Math Program Use and SGP Score. Using a linear regression model, our analysis indicates no positive relationship that was statistically significant between program use in any quartile and SGP scores ($p<.05$). See Appendix A, Table 20 for further details. As illustrated in Figure 12, students who used the ST Math program 31 minutes or less per week (Quartile 1 and 2) had respectively 4.1 and 2.8 lower SGP scores that was statistically significant ($p<.05$), while program users in excess of 31 minutes per week (Quartile 3 & 4) had no statistically significant difference and .6 slightly lower SGP score, as compared to students from comparison schools who did not use the ST Math through a STEM AC license. In summary, in the case of ST Math, regardless of how much time per week spent on the ST math program use, program users in any quartile did not have statistically significant and higher SGP scores as compared to the students from comparison schools who did not use the ST Math program through a STEM AC license.

In addition, other statistically significant and positive relationships with SGP scores were found ($p<.05$) for:

- Asian students who had 3.1 higher SGP scores than White students.
- Rural school students who had 1.9 higher SGP scores than non-rural school students.
- English learner who had 1.5 higher SGP scores than native English speakers.
- 6th graders who had 1.2 higher SGP scores than 4th graders.

See Table 20 in Appendix A for a full list of relationships between student characteristics and SGP score for ST Math.

Discussion

In this section, we provide an overview of findings across the programs, and considerations for implementation improvement and future evaluation and research. Specifically, we illustrate the relationships between weekly program use and RISE math achievement and SGP by providing an overview of results.

We analyzed a total of 56 relationships between weekly program use and student outcomes. Specifically, we examined eight relationships (four for RISE math proficiency with logistical regressions and the same amount for SGP with linear regressions) for each of seven software vendors. This analysis is summarized in Table 6. As noted, statistically significant and positive relationships between any use quartile(s) and any math outcome(s) account for 30% of all 56 relationships while statistically significant and negative relationships constitute 32%. Non-significant relationships represent 38% of our results, which is our most common finding (see Figure 13). Again, statistically significant relationship means that the results are not due to chance ($p<.05$) and are evidence that the intervention—use of the math personalized software program—was related to performance.

Taking all of the findings reported here into consideration, several patterns are worth noting. First, higher rank-ordered use quartiles were related to better math performance. For example, most statistically significant and positive relationships were found in Quartile 4, indicating that students who spent relatively higher amounts of time using the math personalized learning software programs through a STEM AC license performed better than students from comparison schools who did not use them through a STEM AC license.

Results also reveal that lower rank-ordered use quartiles were negatively related to student math outcomes. Specifically, most statistically significant and negatively relationships associated with Quartile 1, illustrating that students who used the math personalized learning software programs through a STEM AC license for relatively little time each week underperformed in comparison to those from comparison schools who did not use them through a STEM AC license.

Finally, middle rank-ordered use quartiles have mixed relationships with student math outcomes. Among students in Quartile 2 and 3, we identified a mix of positive, negative, and non-significant relationships, indicating that there is no clear relationship between moderate levels of software use and math learning outcomes. Based on the findings above, we discuss possible considerations for the math personalized learning software programs below.

Figure 13. Summary of Relationship Patterns between Use Quartiles and RISE Math Proficiency & SGP

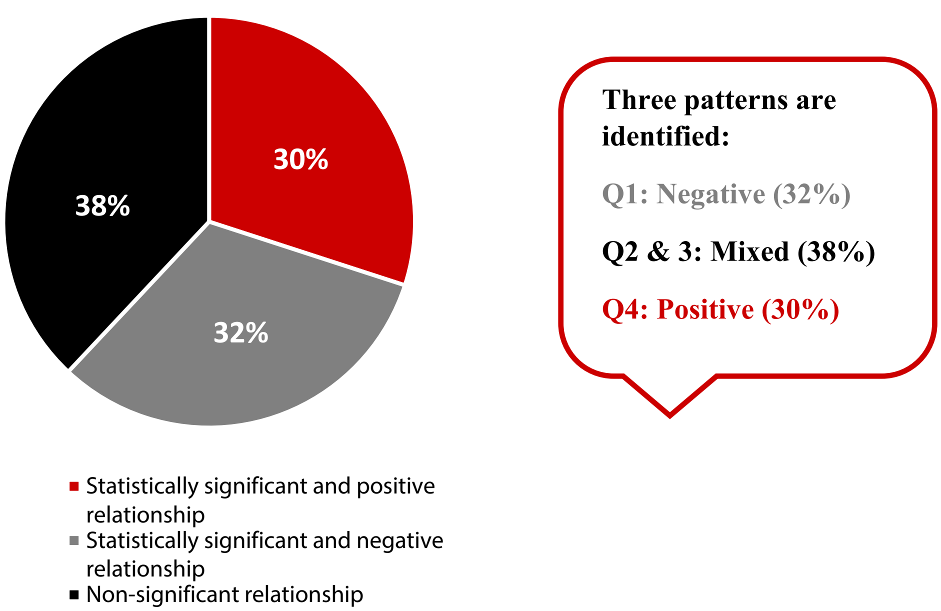


Table 6. Summary of Relationship between Weekly Program Use Quartiles and RISE Math Proficiency & SGP

Vendor	Logistic Regression for RISE Math Proficiency				Linear Regression for SGP			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
ALEKS	−	ns (negative)	+	+	−	+	+	+
DreamBox	ns (positive)	+	ns (positive)	+	ns (positive)	+	ns (positive)	+
Imagine Math	−	−	ns (positive)	+	−	−	ns (negative)	+
i-Ready	−	−	ns (negative)	+	−	−	−	+
Mathspace	ns (negative)	ns (negative)	ns (positive)	+	−	ns (negative)	ns (negative)	+
MobyMax	−	ns (negative)	ns (positive)	ns (positive)	−	ns (positive)	+	+
ST Math	−	−	ns (negative)	ns (negative)	−	−	ns (negative)	ns (negative)

ns (negative)	Non-significant and negative relationship
ns (positive)	Non-significant and positive relationship
−	Statistically significant and negative relationship
+	Statistically significant and positive relationship

Considerations for Implementation Improvement

Encourage additional supplemental use of math personalized learning software programs weekly.

Based on our finding that students who used the software program for relatively longer amounts of time each week performed better than matched students from comparison schools, students should be encouraged to use personalized learning math software programs for a minimum amount of a time each week. This recommendation is consistent with our previous evaluation findings that identified an association between greater program use and better math performance (Owens et al, 2020). This finding is aligned with previous research. For example, Cheung and Slavin (2013) found that a minimum of 30 minutes of weekly math software use optimized student learning outcomes while Heinrich et al. (2019) determined that students who spent more time in their online courses and completed more activities per day had better course performance. Due to the nature of this evaluation, an optimal level of minutes of use cannot be concluded. However, our findings suggest that the most statistically significant and positive relationships between better math performance and software use were found among students with the longest time duration of program use per week, even after controlling for student characteristics. Thus, providing avenues for students to increase their weekly software use may be an effective approach to improve their math outcomes.

Effective integration of math personalized learning software programs into the classroom may improve student engagement and ultimately, math learning outcomes.

Our findings suggest that students who used software programs for relatively little time each week underperformed compared to those from comparison schools. Based on the statistically significant and negative relationships identified, we recommend that exploration of ways to more effectively integrate math software into the classroom, which may improve student engagement with education technology and ultimately, their learning outcomes. As our literature review illustrated, teachers' pedagogical knowledge for appropriately integrating technology into curricula is critical to student learning (Hew & Brush, 2007; Hughes, 2005; Lawless & Pellegrino, 2007; Sarker et al., 2019). A recent empirical study further substantiates the importance of appropriate technology integration, showing that student achievement increased when teachers adjusted their technology use to better align with curriculum and instruction (Callaghan et al., 2018). Another benefit of effective technology integration is to foster student engagement and motivation for learning (Collins & Halverson, 2009). Although it is too early to evaluate, there is a possibility that use of math learning software can become increasingly beneficial when students are in remote learning situations or in discontinuous, or disrupted, learning times such as a schooling during a pandemic. Thus, providing support effective integration of the software program into classroom instruction and student learning may be another approach to improve student engagement and math outcomes, especially for those students who traditionally spent less time on program use than others.

Future Evaluation and Research

Beyond investigating the relationship between time spent using math software and student outcomes, further studies can be conducted on how other patterns of use impact student math outcomes. Given our finding that some program use quartiles had mixed relationships with math outcomes, future research may investigate other student behavioral patterns associated with math performance beyond duration of program use, such as frequency of program use, section and course duration, number of courses enrolled, activities completed per day, idle-to-active time per session, and the percentage of sessions that occur outside the regular school day (Heinrich et al., 2019). Future evaluation of other program use patterns can help us identify other factors that contribute to student math outcomes and allow us to seek other strategies to help students achieve math learning success.

Also, further research can explore how implementation may vary based on student needs. In the current study, our findings of higher program use associated with better math performance are based on sample that is majority White, native English speakers, and non-rural school students. We don't investigate the combination effects of student characteristics and program use on student outcomes due to the primary evaluation focus in the current report, though we provide some findings about relationships between student characteristics and student outcomes as a secondary interest. With a broader goal of educational equity in mind, future study on the impacts of use given student characteristics, such as race/ethnicity, English Learner status, and socioeconomic status would be appropriate to further understand whether any variation in use is related to math performance. In doing so, we may help educators identify and accommodate the unique needs of all students.

Reference List

Adelson, J. L. (2013). Educational research with real-world data: Reducing selection bias with propensity score analysis. *Practical Assessment, Research, and Evaluation*, *18*(1), 15.

Al-Adwan, A., & Smedley, J. (2012). Implementing e-learning in the Jordanian Higher Education System: Factors affecting impact. *International Journal of Education and Development using ICT*, *8*(1).

Auletto, A, Barton, A., Owens, R., & Rorrer, A. K. (2020). *Learning Mathematics with Technology in Utah: An Evaluation of Student Attitudes and Perceptions of Math Personalized Learning Software.* Utah Education Policy Center: Salt Lake City, UT.

Bauer, J. & Kenton, J. (2005). Toward technology integration in the schools why it isn't happening, *Journal of Technology and Teacher Education, 13*(4), 519-547.

Callaghan, M. N., Long, J. J., van Es, E. A., Reich, S. M., & Rutherford, T. (2018). How teachers integrate a math computer game: Professional development use, teaching practices, and student achievement. *Journal of Computer Assisted Learning, 34*(1), 10-19. https://doi.org/10.1111/jcal.12209

Cheung, A.C.K. & Slavin, R.E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9,* 88-113, 10.1016/j.edurev.2013.01.001

Collins, A., & Halverson, R. (2009). *Rethinking education in the age of technology: The Digital revolution and schooling in America*. New York, NY: Teachers College Press.

De Witte, K., Haelermans, C., & Rogge, N. (2015). The effectiveness of a computer-assisted math learning program. *Journal of Computer Assisted Learning, 31,* 314-329. https://doi.org/10.1111/jcal.12090

Eddings, S., McCormick, K., Owens, R., & Rorrer, A.K. (2019) *Utah STEM Action Center Program Evaluation: Academic Year 2018-19.* Utah Education Policy Center: Salt Lake City, UT.

Ertmer, P. (2005). Teacher pedagogical beliefs: the final frontier in our quest for technology ıntegration, *Educational Technology Research and Development, 53*, 25-39. http://dx.doi.org/10.1007/BF02504683

Franklin, J.D., Orians, E., Nicholson, A. & Rorrer, A.K. (2016). *Technology in the Classroom: Strategies, Potential Benefits, & Challenges of Device Adoption*. Utah Education Policy Center. Salt Lake City, UT. https://daqy2hvnfszx3.cloudfront.net/wp-content/uploads/sites/2/2019/02/26091452/Technology-in-the-Classroom.pdf

Heinrich, C. J., Darling-Aduana, J., Good, A., & Cheng, H. (2019). A look inside online educational settings in high school: Promise and pitfalls for improving educational opportunities and outcomes. *American Educational Research Journal*, *56*(6), 2147-2188.

Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational technology research and development*, *55* (3), 223-252.

Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2010). US Math Performance in Global Perspective: How Well Does Each State Do at Producing High-Achieving Students? PEPG Report No.: 10-19. *Program on Education Policy and Governance, Harvard University*.

Higgins K, Huscroft-D'Angelo J, Crawford L. (2019). Effects of technology in mathematics on achievement, motivation, and attitude: A meta-analysis. *Journal of Educational Computing Research, 57*(2), 283-319. https://doi.org/10.1177/0735633117748416

Hughes, J. (2005). The role of teacher knowledge and learning experiences in forming technology-integrated pedagogy. *Journal of technology and teacher education*, *13*(2), 277-302.

Karagiorgi, Y. & Charalambous, K. (2004). Curricula considerations in ICT integration: models and practices in Cyprus, *Education and Information Technologies, 9*(1), 21-35. http://dx.doi.org/10.1023/B:EAIT.0000024259.91951.44

Koedinger, K. R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, *43*(4), 489–510.

Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of educational research*, *77*(4), 575-614.

Li, Q. & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review, 22*(3), 215-243. https://doi.org/10.1007/s10648-010-9125-8

Long, J. S., & Mustillo, S. A. (2018). Using predictions and marginal effects to compare groups in regression models for binary outcomes. *Sociological Methods & Research*, 0049124118799374.

Mishra, P. & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record 108*(6), 1017-1054. http://one2oneheights.pbworks.com/f/MISHRA_PUNYA.pdf

Morgan, G. L. (2008). Improving student engagement: Use of the interactive whiteboard as an instructional tool to improve engagement and behavior in the junior high school classroom. Retrieved from: https://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1140&context=doctoral

Murphy, D. (2016). A literature review: The effect of implementing technology in a high school mathematics classroom. *International Journal of Research in Education and Science (IJRES), 2*(2), 295-299. https://files.eric.ed.gov/fulltext/EJ1105104.pdf

Onuma, F. J., Rorrer, A. K., Pecsok, M., & Weissinger, K. (2020). *Teaching Mathematics with Technology in Utah: An Evaluation of Teacher Knowledge, Practices, and Outcomes with Using Mathematics Personalized Learning Software.* Utah Education Policy Center: Salt Lake City, UT.

Owens, R., Rorrer, A., Ni, Y., Onuma, F. J., Pecsok, M., & Moore, B. (2020). *Longitudinal Evaluation of the Math Personalized Learning Software Grant Program.* Salt Lake City, UT: Utah Education Policy Center.

Pilli, O., & Aksu, M. (2013). The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus. *Computers and Education*, *62*, 62–71.

Puentedura, R. R. (2013). SAMR: Moving from enhancement to transformation. Retrieved from http://www.hippasus.com/rrpweblog/archives/2013/05/29/SAMREnhancementToTransformation.pdf

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., & Gallagher, L. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal*, *47*(4), 833–878.

Sarker, M. N. I., Wu, M., Cao, Q., Alam, G. M., & Li, D. (2019). Leveraging Digital Technology for Better Learning and Education: A Systematic Literature Review. *Int. J. Inf. Educ. Technol*, *9*(7), 453-461.

Schaffhauser, D. (2020). Updated: Free math resources for schools during the COVID-19 outbreak. *The Journal.* https://thejournal.com/articles/2020/04/28/updated-free-math-resources-for-schools-during-the-covid-19-outbreak.aspx

Sharp, L. A., & Hamil, M. (2018). Impact of a Web-based adaptive supplemental digital resource on student mathematics performance. *Online Learning, 22*(1), 81–92. https://eric.ed.gov/?id=EJ1179657

Tatar, E. (2013). The Effect of Dynamic Software on Prospective Mathematics Teachers' Perceptions Regarding Information and Communication Technology. *Australian*

Journal of Teacher Education, 38(12). http://dx.doi.org/10.14221/ajte.2013v38n12.6

Young, J., Gorumek, F. & Hamilton, C. (2018). Technology effectiveness in the mathematics classroom: A systematic review of meta-analytic research. *The Journal of Computer Education, 5*(2), 133-148. https://www.researchgate.net/publication/325105360_

Appendix A. Full Regression Model Results for Each Vendor

ALEKS Regression Model Results

Table 7. Logistic Regression Predicting RISE Math Test Proficiency for ALEKS

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-4.16*	.04	-116.66	<.001	(-4.23,-4.09)	.02	.02	(.01,.02)
Use Quartile								
Quartile 1	-.29*	.03	-10.58	<.001	(-.34,-.23)	.75	.43	(.42,.44)
Quartile 2	-.01	.03	-.32	.747	(-.06,.04)	.99	.50	(.49,.51)
Quartile 3	.16*	.03	6.26	<.001	(.11,.21)	1.17	.54	(.53,.55)
Quartile 4	.27*	.03	10.64	<.001	(.22,.32)	1.31	.57	(.56,.58)
Race/Ethnicity								
Asian	.29*	.06	4.43	<.001	(.16,.41)	1.33	.57	(.54,.6)
African American/Black	-.58*	.09	-6.56	<.001	(-.75,-.4)	.56	.36	(.32,.4)
Latino/Hispanic	-.35*	.03	-13.61	<.001	(-.4,-.3)	.70	.41	(.4,.43)
American Indian/Alaskan Native	-.29*	.09	-3.14	.002	(-.48,-.11)	.75	.43	(.38,.47)
Multiple Races	-.08	.05	-1.66	.097	(-.17,.01)	.93	.48	(.46,.5)
Native Hawaiian/Pacific Islander	-.29*	.07	-4.00	<.001	(-.43,-.15)	.75	.43	(.39,.46)

Table 7. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Grade Level								
3rd	.92	1.08	.86	.392	(-1.19,3.03)	2.51	.72	(.23,.95)
5th	-.03	.03	-1.08	.28	(-.1,.03)	.97	.49	(.48,.51)
6th	-.86*	.03	-27.87	<.001	(-.92,-.8)	.42	.30	(.28,.31)
7th	-.02	.03	-.64	.525	(-.08,.04)	.98	.50	(.48,.51)
8th	-.03	.03	-.89	.374	(-.09,.03)	.97	.49	(.48,.51)
9th	-.58*	.03	-19.26	<.001	(-.64,-.52)	.56	.36	(.34,.37)
10th	-1.04*	.03	-33.20	<.001	(-1.1,-.98)	.35	.26	(.25,.27)
11th	-.99*	.42	-2.34	.019	(-1.82,-.16)	.37	.27	(.14,.46)
Other Student Characteristics								
Previous Year Proficient	1.78*	.01	191.43	<.001	(1.76,1.79)	5.91	.86	(.85,.86)
Male	.01	.02	.72	.473	(-.02,.04)	1.01	.50	(.5,.51)
Rural	.10*	.02	4.64	<.001	(.06,.15)	1.11	.53	(.52,.54)
Chronic Absence	-.60*	.03	-21.43	<.001	(-.65,-.54)	.55	.35	(.34,.37)
Special Education	-.65*	.03	-20.21	<.001	(-.71,-.59)	.52	.34	(.33,.36)
English Learner	-.23*	.04	-5.30	<.001	(-.32,-.15)	.79	.44	(.42,.46)
Free and Reduced Lunch	-.31*	.02	-17.93	<.001	(-.35,-.28)	.73	.42	(.41,.43)
Mobile	-.54*	.05	-11.24	<.001	(-.63,-.44)	.58	.37	(.35,.39)

*Indicates statistical significance at the p<.05 level

Source: Student Records and Vendor Data

Table 8. Linear Regression Predicting SGP Score for ALEKS

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	57.07*	.34	167.91	<.001	(56.41,57.74)
Use Quartile					
Quartile 1	-3.06*	.27	-11.22	<.001	(-3.6,-2.53)
Quartile 2	.84*	.27	3.16	.002	(.32,1.37)
Quartile 3	3.31*	.27	12.46	<.001	(2.79,3.83)
Quartile 4	5.19*	.27	19.01	<.001	(4.66,5.73)
Race/Ethnicity					
Asian	3.45*	.69	5.00	<.001	(2.1,4.8)
African American/Black	-4.57*	.75	-6.06	<.001	(-6.04,-3.09)
Latino/Hispanic	-2.85*	.26	-11.06	<.001	(-3.35,-2.34)
American Indian/Alaskan Native	-2.44*	.82	-2.98	.003	(-4.04,-.83)
Multiple Races	-.20	.49	-.42	.678	(-1.17,.76)
Native Hawaiian/Pacific Islander	-2.65*	.70	-3.78	<.001	(-4.02,-1.27)
Grade Level					
5^{th}	-1.15*	.32	-3.62	<.001	(-1.77,-.53)
6^{th}	-1.49*	.31	-4.83	<.001	(-2.09,-.89)
7^{th}	-1.67*	.30	-5.54	<.001	(-2.26,-1.08)
8^{th}	-2.70*	.31	-8.76	<.001	(-3.3,-2.09)
9^{th}	-3.59*	.31	-11.48	<.001	(-4.21,-2.98)
10^{th}	-5.42*	.32	-16.86	<.001	(-6.05,-4.79)
11^{th}	-47.56*	2.79	-17.05	<.001	(-53.02,-42.09)
Other Student Characteristics					
Previous Year Proficient	-1.13*	.08	-14.68	<.001	(-1.28,-.98)
Male	-.80*	.16	-5.14	<.001	(-1.11,-.5)
Rural	1.35*	.24	5.72	<.001	(.89,1.81)
Chronic Absence	-5.55*	.27	-20.56	<.001	(-6.08,-5.02)
Special Education	-4.63*	.25	-18.21	<.001	(-5.13,-4.13)
English Learner	.34	.36	.94	.347	(-.37,1.05)
Free and Reduced Lunch	-3.00*	.18	-16.63	<.001	(-3.36,-2.65)
Mobile	-4.02*	.45	-8.90	<.001	(-4.9,-3.13)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

DreamBox Regression Model Results

Table 9. Logistic Regression Predicting RISE Math Test for DreamBox

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-4.08*	.30	-13.77	<.001	(-4.67,-3.5)	.02	.02	(.01,.03)
Use Quartile								
Quartile 1	.23	.27	.88	.379	(-.29,.76)	1.26	.56	(.43,.68)
Quartile 2	.84*	.28	3.05	.002	(.3,1.38)	2.31	.70	(.57,.8)
Quartile 3	.31	.27	1.15	.251	(-.22,.83)	1.36	.58	(.45,.7)
Quartile 4	1.39*	.35	3.98	<.001	(.7,2.07)	4.01	.80	(.67,.89)
Race/Ethnicity								
Asian	.25	.73	.35	.727	(-1.17,1.68)	1.29	.56	(.24,.84)
African American/Black	-1.04	.88	-1.19	.234	(-2.76,.67)	.35	.26	(.06,.66)
Latino/Hispanic	-.53*	.26	-2.02	.044	(-1.05,-.01)	.59	.37	(.26,.5)
American Indian/Alaskan Native	-.75	1.04	-.72	.47	(-2.79,1.29)	.47	.32	(.06,.78)
Multiple Races	-.62	.36	-1.72	.085	(-1.32,.09)	.54	.35	(.21,.52)
Native Hawaiian/Pacific Islander	-.53	.93	-.57	.569	(-2.36,1.3)	.59	.37	(.09,.79)
Grade Level								
5th	.09	.17	.51	.608	(-.25,.43)	1.09	.52	(.44,.61)
6th	.12	.25	.47	.636	(-.38,.62)	1.13	.53	(.41,.65)

Table 9. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Other Student Characteristics								
Previous Year Proficient	1.60*	.09	17.88	<.001	(1.42,1.78)	4.95	.83	(.81,.86)
Male	.29	.16	1.81	.071	(-.02,.61)	1.34	.57	(.49,.65)
Chronic Absence	-.08	.25	-.31	.756	(-.58,.42)	.92	.48	(.36,.6)
Special Education	-.11	.24	-.46	.647	(-.58,.36)	.90	.47	(.36,.59)
English Learner	.66	.40	1.66	.096	(-.12,1.45)	1.94	.66	(.47,.81)
Free and Reduced Lunch	-.13	.18	-.72	.471	(-.48,.22)	.88	.47	(.38,.56)
Mobile	-1.40*	.52	-2.68	.007	(-2.42,-.37)	.25	.20	(.08,.41)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

Table 10. Linear Regression Predicting SGP Score for DreamBox

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	60.66	2.86	21.20	<.001	(55.05,66.27)
Use Quartile					
Quartile 1	2.92	2.80	1.05	.296	(-2.56,8.41)
Quartile 2	12.86*	2.84	4.53	<.001	(7.29,18.43)
Quartile 3	4.68	2.93	1.60	.111	(-1.07,10.44)
Quartile 4	10.15*	3.16	3.21	.001	(3.95,16.35)
Race/Ethnicity					
Asian	-4.05	7.14	-.57	.571	(-18.05,9.95)
African American/Black	-18.00*	8.85	-2.03	.042	(-35.35,-.64)
Latino/Hispanic	-3.52	2.79	-1.26	.208	(-9,1.96)
American Indian/Alaskan Native	-13.05	10.52	-1.24	.215	(-33.68,7.57)
Multiple Races	1.63	3.77	.43	.666	(-5.76,9.02)
Native Hawaiian/Pacific Islander	17.30	8.95	1.93	.054	(-.24,34.85)
Grade Level					
5^{th}	.80	1.82	.44	.661	(-2.77,4.37)
6^{th}	14.74*	2.70	5.47	<.001	(9.46,20.03)
Other Student Characteristics					
Previous Year Proficient	-2.67*	.78	-3.41	.001	(-4.21,-1.14)
Male	1.73	1.69	1.02	.307	(-1.58,5.03)
Chronic Absence	-4.66	2.72	-1.71	.087	(-9.98,.67)
Special Education	-6.65*	2.38	-2.79	.005	(-11.32,-1.99)
English Learner	2.77	4.20	.66	.511	(-5.47,11)
Free and Reduced Lunch	-2.14	1.93	-1.11	.268	(-5.92,1.64)
Mobile	-20.14*	5.21	-3.86	<.001	(-30.36,-9.93)

*Indicates statistical significance at the p<.05 level

Source: Student Records and Vendor Data

Imagine Math Regression Model Results

Table 11. Logistic Regression Predicting RISE Math Test Proficiency for Imagine Math

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-4.07*	.05	-74.54	<.001	(-4.17,-3.96)	.02	.02	(.02,.02)
Use Quartile								
Quartile 1	-.29*	.05	-5.73	<.001	(-.39,-.19)	.75	.43	(.4,.45)
Quartile 2	-.16*	.05	-3.21	.001	(-.26,-.06)	.85	.46	(.44,.48)
Quartile 3	.03	.05	.52	.603	(-.07,.12)	1.03	.51	(.48,.53)
Quartile 4	.30*	.05	6.16	<.001	(.2,.39)	1.35	.57	(.55,.6)
Race/Ethnicity								
Asian	.03	.11	.27	.789	(-.19,.25)	1.03	.51	(.45,.56)
African American/Black	-.39*	.15	-2.66	.008	(-.68,-.1)	.68	.40	(.34,.47)
Latino/Hispanic	-.22*	.05	-4.57	<.001	(-.32,-.13)	.80	.44	(.42,.47)
American Indian/Alaskan Native	-.42	.22	-1.93	.054	(-.84,.01)	.66	.40	(.3,.5)
Multiple Races	-.22*	.08	-2.96	.003	(-.37,-.08)	.80	.44	(.41,.48)
Native Hawaiian/Pacific Islander	-.11	.13	-.81	.416	(-.36,.15)	.90	.47	(.41,.54)
Grade Level								
3rd	1.30	1.21	1.07	.286	(-1.08,3.67)	3.65	.79	(.25,.98)
5th	.02	.04	.54	.589	(-.05,.09)	1.02	.51	(.49,.52)
6th	-.62*	.04	-16.19	<.001	(-.69,-.54)	.54	.35	(.33,.37)
7th	.06	.05	1.09	.276	(-.05,.17)	1.06	.51	(.49,.54)
8th	.32*	.06	5.14	<.001	(.2,.44)	1.38	.58	(.55,.61)
9th	-.18*	.07	-2.44	.015	(-.32,-.03)	.84	.46	(.42,.49)
10th	-.63*	.07	-9.31	<.001	(-.76,-.5)	.53	.35	(.32,.38)
11th	-.12	1.29	-.09	.925	(-2.64,2.4)	.89	.47	(.07,.92)

Table 11. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Other Student Characteristics								
Previous Year Proficient	1.67*	.02	105.51	<.001	(1.64,1.7)	5.32	.84	(.84,.85)
Male	.12*	.03	4.61	<.001	(.07,.18)	1.13	.53	(.52,.54)
Rural	.15*	.05	3.03	.002	(.05,.24)	1.16	.54	(.51,.56)
Chronic Absence	-.45*	.05	-9.94	<.001	(-.54,-.36)	.64	.39	(.37,.41)
Special Education	-.63*	.05	-13.19	<.001	(-.73,-.54)	.53	.35	(.33,.37)
English Learner	-.02	.07	-.34	.734	(-.16,.11)	.98	.49	(.46,.53)
Free and Reduced Lunch	-.26*	.03	-7.74	<.001	(-.32,-.19)	.77	.44	(.42,.45)
Mobile	-.33*	.09	-3.81	<.001	(-.5,-.16)	.72	.42	(.38,.46)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

Table 12. Linear Regression Predicting SGP Score for Imagine Math

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	55.84*	.53	106.06	<.001	(54.8,56.87)
Use Quartile					
Quartile 1	-5.37*	.53	-10.20	<.001	(-6.4,-4.33)
Quartile 2	-3.44*	.52	-6.58	<.001	(-4.47,-2.42)
Quartile 3	-.59	.52	-1.15	.249	(-1.61,.42)
Quartile 4	3.64*	.50	7.27	<.001	(2.66,4.62)
Race/Ethnicity					
Asian	1.77	1.23	1.44	.15	(-.64,4.17)
African American/Black	-4.45*	1.36	-3.28	.001	(-7.11,-1.8)
Latino/Hispanic	-1.92*	.51	-3.78	<.001	(-2.92,-.92)
American Indian/Alaskan Native	-1.59	2.02	-.79	.432	(-5.56,2.37)
Multiple Races	-2.03*	.81	-2.52	.012	(-3.62,-.45)
Native Hawaiian/Pacific Islander	-3.26*	1.29	-2.54	.011	(-5.78,-.74)
Grade Level					
5th	-.07	.38	-.20	.845	(-.82,.67)
6th	2.20*	.39	5.63	<.001	(1.44,2.97)
7th	2.15*	.57	3.76	<.001	(1.03,3.27)
8th	5.32*	.70	7.60	<.001	(3.95,6.69)
9th	-.41	.80	-.52	.605	(-1.98,1.16)
10th	-2.98*	.80	-3.73	<.001	(-4.55,-1.42)
11th	-44.42*	11.76	-3.78	<.001	(-67.47,-21.37)
Other Student Characteristics					
Previous Year Proficient	-1.10*	.14	-8.12	<.001	(-1.37,-.84)
Male	.09	.28	.30	.761	(-.47,.64)
Rural	2.51*	.50	4.99	<.001	(1.53,3.5)
Chronic Absence	-4.54*	.46	-9.88	<.001	(-5.44,-3.64)
Special Education	-6.76*	.43	-15.63	<.001	(-7.61,-5.91)
English Learner	2.01*	.67	2.99	.003	(.69,3.34)
Free and Reduced Lunch	-2.78*	.35	-7.98	<.001	(-3.46,-2.1)
Mobile	-4.62*	.86	-5.37	<.001	(-6.31,-2.94)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

i-Ready Regression Model Results

Table 13. Logistic Regression Predicting RISE Math Test Proficiency for i-Ready

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-3.90*	.05	-75.62	<.001	(-4.00,-3.79)	.02	.02	(.02,.02)
Use Quartile								
Quartile 1	-.56*	.05	-11.55	<.001	(-0.66,-0.47)	.57	.36	(.34,.39)
Quartile 2	-.27*	.05	-5.64	<.001	(-0.36,-0.17)	.77	.43	(.41,.46)
Quartile 3	-.02	.05	-.42	.672	(-0.12,0.08)	.98	.49	(.47,.52)
Quartile 4	.12*	.05	2.42	.016	(0.02,0.22)	1.13	.53	(.51,.56)
Race/Ethnicity								
Asian	.30*	.12	2.55	.011	(0.07,0.53)	1.35	.57	(.52,.63)
African American/Black	-.57*	.14	-4.12	<.001	(-0.85,-0.30)	.56	.36	(.3,.43)
Latino/Hispanic	-.30*	.04	-6.84	<.001	(-0.39,-0.22)	.74	.42	(.4,.45)
American Indian/Alaskan Native	-.34	.19	-1.76	.079	(-0.71,0.04)	.71	.42	(.33,.51)
Multiple Races	-.04	.07	-.52	.602	(-0.18,0.11)	.96	.49	(.45,.53)
Native Hawaiian/Pacific Islander	-.21	.13	-1.68	.094	(-0.46,0.04)	.81	.45	(.39,.51)
Grade Level								
3rd	1.45	1.16	1.25	.211	(-0.82,3.71)	4.25	.81	(.31,.98)
5th	.00	.04	-.06	.95	(-0.07,0.07)	1.00	.50	(.48,.52)
6th	-.82*	.04	-21.39	<.001	(-0.89,-0.74)	.44	.31	(.29,.32)
7th	-.14*	.05	-2.85	.004	(-0.24,-0.04)	.87	.46	(.44,.49)
8th	-.38*	.06	-6.72	<.001	(-0.49,-0.27)	.68	.41	(.38,.43)
9th	-.98*	.07	-15.03	<.001	(-1.10,-0.85)	.38	.27	(.25,.3)
10th	-1.24*	.06	-20.41	<.001	(-1.36,-1.12)	.29	.22	(.2,.25)
11th	-.22	.64	-.34	.736	(-1.47,1.04)	.81	.45	(.19,.74)

Table 13. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Other Student Characteristics								
Previous Year Proficient	1.69*	.02	111.42	<.001	(1.66,1.72)	5.43	.84	(.84,.85)
Male	.07*	.03	2.62	.009	(0.02,0.12)	1.07	.52	(.5,.53)
Rural	.06	.05	1.42	.155	(-0.02,0.15)	1.07	.52	(.49,.54)
Chronic Absence	-.49*	.04	-11.54	<.001	(-0.57,-0.41)	.61	.38	(.36,.4)
Special Education	-.61*	.05	-12.70	<.001	(-0.70,-0.51)	.54	.35	(.33,.37)
English Learner	-.13*	.06	-2.22	.026	(-0.25,-0.02)	.87	.47	(.44,.5)
Free and Reduced Lunch	-.27*	.03	-9.08	<.001	(-0.33,-0.21)	.76	.43	(.42,.45)
Mobile	-.54*	.08	-6.94	<.001	(-0.69,-0.39)	.58	.37	(.33,.4)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

Table 14. Linear Regression Predicting SGP Score for i-Ready

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	58.73*	.49	118.70	<.001	(57.76,59.7)
Use Quartile					
Quartile 1	-6.21*	.48	-13.00	<.001	(-7.15,-5.27)
Quartile 2	-3.41*	.48	-7.11	<.001	(-4.34,-2.47)
Quartile 3	-2.23*	.50	-4.47	<.001	(-3.21,-1.25)
Quartile 4	2.16*	.50	4.32	<.001	(1.18,3.14)
Race/Ethnicity					
Asian	3.95*	1.23	3.23	.001	(1.55,6.36)
African American/Black	-4.49*	1.21	-3.72	<.001	(-6.85,-2.12)
Latino/Hispanic	-2.68*	.44	-6.06	<.001	(-3.55,-1.81)
American Indian/Alaskan Native	-3.91*	1.68	-2.33	.02	(-7.2,-.63)
Multiple Races	.78	.76	1.02	.306	(-0.72,2.28)
Native Hawaiian/Pacific Islander	-.64	1.22	-.52	.601	(-3.02,1.75)
Grade Level					
5th	-.61	.36	-1.69	.091	(-1.31,.1)
6th	-.79*	.38	-2.06	.04	(-1.53,-.04)
7th	-1.97*	.50	-3.99	<.001	(-2.94,-1)
8th	-7.09*	.57	-12.35	<.001	(-8.21,-5.96)
9th	-9.22*	.73	-12.64	<.001	(-10.65,-7.79)
10th	-14.33*	.68	-21.08	<.001	(-15.66,-13)
11th	-48.76*	5.85	-8.33	<.001	(-60.23,-37.29)
Other Student Characteristics					
Previous Year Proficient	-1.15*	.13	-9.03	<.001	(-1.4,-.9)
Male	-.77*	.26	-2.90	.004	(-1.29,-.25)
Rural	.29	.46	.62	.536	(-0.62,1.19)
Chronic Absence	-4.32*	.41	-10.41	<.001	(-5.13,-3.5)
Special Education	-5.51*	.40	-13.67	<.001	(-6.3,-4.72)
English Learner	1.27*	.55	2.33	.02	(0.2,2.35)
Free and Reduced Lunch	-2.97*	.31	-9.67	<.001	(-3.58,-2.37)
Mobile	-4.15*	.73	-5.65	<.001	(-5.59,-2.71)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

Mathspace Regression Model Results

Table 15. Logistic Regression Predicting RISE Math Test Proficiency for Mathspace

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-3.96*	.11	-37.47	<.001	(-4.16,-3.75)	.02	.02	(.02,.02)
Use Quartile								
Quartile 1	-.20	.13	-1.50	.133	(-0.45,0.06)	.82	.45	(.39,.51)
Quartile 2	-.05	.13	-.40	.692	(-0.31,0.21)	.95	.49	(.42,.55)
Quartile 3	.14	.13	1.08	.282	(-0.12,0.40)	1.15	.54	(.47,.6)
Quartile 4	.37*	.13	2.95	.003	(0.13,0.62)	1.45	.59	(.53,.65)
Race/Ethnicity								
Asian	-.09	.26	-.36	.72	(-0.61,0.42)	.91	.48	(.35,.6)
African American/Black	-.53	.35	-1.51	.132	(-1.23,0.16)	.59	.37	(.23,.54)
Latino/Hispanic	-.15	.09	-1.54	.123	(-0.33,0.04)	.86	.46	(.42,.51)
American Indian/Alaskan Native	-.27	.31	-.87	.382	(-0.88,0.34)	.76	.43	(.29,.58)
Multiple Races	-.27	.16	-1.72	.086	(-0.57,0.04)	.76	.43	(.36,.51)
Native Hawaiian/Pacific Islander	.12	.26	.47	.64	(-0.39,0.63)	1.13	.53	(.4,.65)
Grade Level								
3rd	-7.79	119.47	-.07	.948	(-241.95,226.37)	.00	.00	(0,1)
5th	-.09	.08	-1.12	.263	(-0.26,0.07)	.91	.48	(.44,.52)
6th	-1.08*	.08	-13.40	<.001	(-1.23,-0.92)	.34	.25	(.23,.29)
7th	-.16	.09	-1.78	.076	(-0.35,0.02)	.85	.46	(.41,.5)
8th	-.20*	.10	-2.01	.044	(-0.39,-0.01)	.82	.45	(.4,.5)
9th	-.78*	.10	-7.70	<.001	(-0.98,-0.58)	.46	.31	(.27,.36)
10th	-1.00*	.10	-10.24	<.001	(-1.20,-0.81)	.37	.27	(.23,.31)
11th	.19	.95	.20	.84	(-1.68,2.06)	1.21	.55	(.16,.89)

Table 15. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Other Student Characteristics								
Previous Year Proficient	1.72*	.03	57.41	<.001	(1.67,1.78)	5.61	.85	(.84,.86)
Male	.01	.05	.24	.812	(-0.09,0.11)	1.01	.50	(.48,.53)
Rural	.32*	.07	4.82	<.001	(0.19,0.45)	1.37	.58	(.55,.61)
Chronic Absence	-.40*	.08	-4.81	<.001	(-0.57,-0.24)	.67	.40	(.36,.44)
Special Education	-.66*	.09	-7.00	<.001	(-0.84,-0.48)	.52	.34	(.3,.38)
English Learner	-.33*	.14	-2.43	.015	(-0.60,-0.06)	.72	.42	(.35,.48)
Free and Reduced Lunch	-.28*	.06	-5.07	<.001	(-0.39,-0.17)	.75	.43	(.4,.46)
Mobile	-.41*	.16	-2.59	.01	(-0.72,-0.10)	.66	.40	(.33,.48)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

Table 16. Linear Regression Predicting SGP Score for Mathspace

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	58.16*	1.00	58.03	<.001	(56.19,60.12)
Use Quartile					
Quartile 1	-4.67*	1.32	-3.55	<.001	(-7.25,-2.09)
Quartile 2	-1.32	1.31	-1.01	.313	(-3.9,1.25)
Quartile 3	-.92	1.30	-.71	.48	(-3.48,1.63)
Quartile 4	3.92*	1.29	3.03	.002	(1.39,6.46)
Race/Ethnicity					
Asian	2.42	2.85	.85	.395	(-3.16,8)
African American/Black	-3.81	2.98	-1.28	.201	(-9.65,2.02)
Latino/Hispanic	-.90	.93	-.97	.333	(-2.73,.92)
American Indian/Alaskan Native	-5.14	2.73	-1.88	.06	(-10.49,.22)
Multiple Races	-.56	1.63	-.34	.732	(-3.75,2.64)
Native Hawaiian/Pacific Islander	-3.95	2.57	-1.54	.124	(-8.98,1.09)
Grade Level					
5th	-1.54	.83	-1.87	.062	(-3.16,.08)
6th	-3.51*	.78	-4.49	<.001	(-5.04,-1.98)
7th	1.17	.93	1.25	.21	(-.66,2.99)
8th	-7.95*	1.00	-7.97	<.001	(-9.91,-6)
9th	-3.50*	1.10	-3.18	.001	(-5.65,-1.34)
10th	-8.28*	1.04	-7.95	<.001	(-10.32,-6.24)
11th	-48.02*	9.04	-5.31	<.001	(-65.75,-30.3)
Other Student Characteristics					
Previous Year Proficient	-1.27*	.25	-5.08	<.001	(-1.76,-.78)
Male	-1.27*	.51	-2.49	.013	(-2.26,-.27)
Rural	5.52*	.68	8.10	<.001	(4.18,6.86)
Chronic Absence	-2.85*	.82	-3.48	<.001	(-4.45,-1.25)
Special Education	-5.29*	.78	-6.82	<.001	(-6.81,-3.77)
English Learner	-1.08	1.21	-.90	.371	(-3.45,1.29)
Free and Reduced Lunch	-2.95*	.57	-5.16	<.001	(-4.07,-1.83)
Mobile	-6.27*	1.52	-4.12	<.001	(-9.25,-3.29)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

MobyMax Regression Model Results

Table 17. Logistic Regression Predicting RISE Math Test Proficiency for MobyMax

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-4.50*	.11	-4.16	<.001	(-4.72,-4.28)	.01	.01	(.01,.01)
Use Quartile								
Quartile 1	-.82*	.28	-2.90	.004	(-1.37,-.27)	.44	.31	(.2,.43)
Quartile 2	-.06	.28	-.21	.836	(-.6,.49)	.94	.49	(.35,.62)
Quartile 3	.26	.27	.95	.34	(-.27,.79)	1.29	.56	(.43,.69)
Quartile 4	.43	.38	1.13	.26	(-.32,1.18)	1.54	.61	(.42,.76)
Race/Ethnicity								
Asian	.19	.23	.82	.414	(-.26,.64)	1.21	.55	(.43,.65)
African American/Black	-.89*	.31	-2.88	.004	(-1.49,-.28)	.41	.29	(.18,.43)
Latino/Hispanic	-.21*	.09	-2.27	.023	(-.39,-.03)	.81	.45	(.4,.49)
American Indian/Alaskan Native	-.18	.44	-.40	.687	(-1.04,.69)	.84	.46	(.26,.67)
Multiple Races	-.39*	.14	-2.72	.007	(-.67,-.11)	.68	.40	(.34,.47)
Native Hawaiian/Pacific Islander	-.45*	.27	-1.67	.094	(-.97,.08)	.64	.39	(.28,.52)
Grade Level								
3rd	2.34	1.56	1.50	.134	(-.72,5.39)	1.33	.91	(.33,1)
5th	.13	.09	1.51	.132	(-.04,.31)	1.14	.53	(.49,.58)
6th	-.46*	.09	-5.42	<.001	(-.63,-.3)	.63	.39	(.35,.43)
7th	.41*	.11	3.73	<.001	(.19,.62)	1.50	.60	(.55,.65)
8th	.20	.11	1.94	.052	(0,.41)	1.23	.55	(.5,.6)
9th	.09	.12	.76	.446	(-.15,.33)	1.10	.52	(.46,.58)
10th	-.95*	.08	-11.82	<.001	(-1.11,-.79)	.39	.28	(.25,.31)
11th	-.94	1.10	-.85	.395	(-3.09,1.22)	.39	.28	(.04,.77)

Table 17. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Other Student Characteristics								
Previous Year Proficient	1.79*	.03	57.72	<.001	(1.73,1.85)	5.97	.86	(.85,.86)
Male	.13*	.05	2.59	.01	(.03,.23)	1.14	.53	(.51,.56)
Rural	-.03	.17	-.16	.875	(-.36,.3)	.97	.49	(.41,.58)
Chronic Absence	-.65*	.11	-6.14	<.001	(-.86,-.45)	.52	.34	(.3,.39)
Special Education	-.64*	.10	-6.44	<.001	(-.83,-.44)	.53	.35	(.3,.39)
English Learner	-.21	.18	-1.17	.243	(-.57,.14)	.81	.45	(.36,.54)
Free and Reduced Lunch	-.10	.07	-1.55	.12	(-.23,.03)	.90	.47	(.44,.51)
Mobile	-.33*	.15	-2.13	.033	(-.63,-.03)	.72	.42	(.35,.49)

*Indicates statistical significance at the p<.05 level

Source: Student Records and Vendor Data

Table 18. Linear Regression Predicting SGP Score for MobyMax

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	55.27*	1.04	52.99	<.001	(53.22,57.31)
Use Quartile					
Quartile 1	-5.47*	2.75	-1.99	.047	(-1.87,-.08)
Quartile 2	1.48	2.75	.54	.591	(-3.91,6.86)
Quartile 3	8.44*	2.54	3.33	.001	(3.47,13.42)
Quartile 4	8.90*	3.28	2.71	.007	(2.47,15.34)
Race/Ethnicity					
Asian	1.86	2.41	.77	.44	(-2.87,6.6)
African American/Black	-5.70*	2.41	-2.36	.018	(-1.42,-.98)
Latino/Hispanic	-.91	.91	-1.00	.317	(-2.68,.87)
American Indian/Alaskan Native	-1.84	4.41	-.42	.677	(-1.48,6.81)
Multiple Races	-4.32*	1.48	-2.91	.004	(-7.23,-1.41)
Native Hawaiian/Pacific Islander	-5.24*	2.41	-2.18	.029	(-9.95,-.52)
Grade Level					
5th	-1.30	.88	-1.47	.14	(-3.02,.43)
6th	3.00*	.86	3.47	.001	(1.31,4.7)
7th	3.14*	1.08	2.91	.004	(1.03,5.25)
8th	5.13*	1.12	4.60	<.001	(2.95,7.32)
9th	6.12*	1.35	4.53	<.001	(3.47,8.77)
10th	-8.58*	.84	-1.22	<.001	(-1.23,-6.94)
11th	-43.79*	6.44	-6.81	<.001	(-56.41,-31.18)
Other Student Characteristics					
Previous Year Proficient	-1.43*	.26	-5.60	<.001	(-1.94,-.93)
Male	-.61	.52	-1.19	.236	(-1.62,.4)
Rural	2.15	1.77	1.22	.224	(-1.32,5.63)
Chronic Absence	-4.78*	1.04	-4.58	<.001	(-6.82,-2.73)
Special Education	-7.75*	.82	-9.42	<.001	(-9.36,-6.14)
English Learner	1.81	1.47	1.23	.219	(-1.08,4.7)
Free and Reduced Lunch	-1.82*	.67	-2.72	.007	(-3.13,-.51)
Mobile	-5.72*	1.49	-3.83	<.001	(-8.65,-2.79)

*Indicates statistical significance at the $p<.05$ level

Source: Student Records and Vendor Data

ST Math Regression Model Results

Table 19. Logistic Regression Predicting RISE Math Test Proficiency for ST Math

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Intercept	-3.96*	.05	-79.18	<.001	(-4.06,-3.87)	.02	.02	(.02,.02)
Use Quartile								
Quartile 1	-.41*	.06	-6.93	<.001	(-.53,-.29)	.66	.40	(.37,.43)
Quartile 2	-.26*	.06	-4.57	<.001	(-.38,-.15)	.77	.43	(.41,.46)
Quartile 3	-.02	.06	-.33	.743	(-.13,.09)	.98	.50	(.47,.52)
Quartile 4	-.01	.06	-.25	.805	(-.12,.1)	.99	.50	(.47,.52)
Race/Ethnicity								
Asian	.30*	.09	3.50	<.001	(.13,.47)	1.35	.58	(.53,.62)
African American/Black	-.69*	.11	-6.38	<.001	(-.91,-.48)	.50	.33	(.29,.38)
Latino/Hispanic	-.36*	.04	-8.64	<.001	(-.44,-.27)	.70	.41	(.39,.43)
American Indian/Alaskan Native	-.33*	.14	-2.27	.023	(-.61,-.04)	.72	.42	(.35,.49)
Multiple Races	-.16*	.08	-2.18	.029	(-.31,-.02)	.85	.46	(.42,.5)
Native Hawaiian/Pacific Islander	-.18*	.09	-1.92	.054	(-.35,0)	.84	.46	(.41,.5)
Grade Level								
3rd	-.57	2.57	-.22	.824	(-5.61,4.46)	.56	.36	(0,.99)
5th	-.01	.03	-.19	.85	(-.07,.06)	.99	.50	(.48,.52)
6th	-.69*	.04	-18.77	<.001	(-.77,-.62)	.50	.33	(.32,.35)
7th	-.06	.05	-1.32	.187	(-.16,.03)	.94	.48	(.46,.51)
8th	-.20*	.05	-4.11	<.001	(-.3,-.11)	.82	.45	(.43,.47)
9th	-.51*	.06	-9.00	<.001	(-.62,-.4)	.60	.38	(.35,.4)
10th	-.82*	.08	-1.45	<.001	(-.98,-.67)	.44	.30	(.27,.34)
11th	-8.65	64.48	-.13	.893	(-135.03,117.74)	.00	.00	(0,1)

Table 19. (Continued)

Predictor Variable	Estimate	Standard Error	*z*	*p*	95% CI for Estimate	Odds	Probability	95% CI for Probability
Student Characteristics								
Previous Year Proficient	1.69*	.01	116.85	<.001	(1.67,1.72)	5.45	.84	(.84,.85)
Male	.08*	.02	3.29	.001	(.03,.13)	1.09	.52	(.51,.53)
Rural	.12*	.04	2.92	.003	(.04,.2)	1.13	.53	(.51,.55)
Chronic Absence	-.44*	.04	-1.35	<.001	(-.52,-.36)	.64	.39	(.37,.41)
Special Education	-.62*	.05	-13.69	<.001	(-.71,-.53)	.54	.35	(.33,.37)
English Learner	-.05	.05	-.92	.357	(-.14,.05)	.96	.49	(.46,.51)
Free and Reduced Lunch	-.25*	.03	-8.50	<.001	(-.31,-.19)	.78	.44	(.42,.45)
Mobile	-.47*	.07	-6.37	<.001	(-.61,-.32)	.63	.39	(.35,.42)

*Indicates statistical significance at the p<.05 level

Source: Student Records and Vendor Data

Table 20. Linear Regression Predicting SGP Score for ST Math

Predictor Variable	Estimate	St Error	*t*	*p*	95% CI
Intercept	57.74*	.48	119.78	<.001	(56.79,58.68)
Use Quartile					
Quartile 1	-4.13*	.55	-7.47	<.001	(-5.22,-3.05)
Quartile 2	-2.75*	.56	-4.90	<.001	(-3.84,-1.65)
Quartile 3	-.61	.56	-1.10	.273	(-1.72,.49)
Quartile 4	-.02	.57	-.03	.976	(-1.13,1.09)
Race/Ethnicity					
Asian	3.07*	.90	3.41	.001	(1.31,4.83)
African American/Black	-5.23*	.94	-5.55	<.001	(-7.07,-3.38)
Latino/Hispanic	-2.43*	.42	-5.80	<.001	(-3.25,-1.61)
American Indian/Alaskan Native	-2.56*	1.30	-1.97	.049	(-5.11,-.01)
Multiple Races	-1.02	.79	-1.29	.197	(-2.57,.53)
Native Hawaiian/Pacific Islander	-2.07*	.91	-2.28	.023	(-3.85,-.29)
Grade Level					
5th	-.66	.34	-1.94	.052	(-1.34,.01)
6th	1.23*	.37	3.35	.001	(.51,1.95)
7th	-1.54*	.50	-3.08	.002	(-2.51,-.56)
8th	-2.59*	.53	-4.92	<.001	(-3.62,-1.56)
9th	-1.40*	.65	-2.14	.032	(-2.67,-.12)
10th	-3.01*	.91	-3.29	.001	(-4.81,-1.22)
11th	-44.24*	16.56	-2.67	.008	(-76.7,-11.79)
Other Student Characteristics					
Previous Year Proficient	-1.25*	.12	-1.13	<.001	(-1.5,-1.01)
Male	-.41	.26	-1.59	.111	(-.91,.09)
Rural	1.91*	.43	4.47	<.001	(1.07,2.75)
Chronic Absence	-5.11*	.41	-12.42	<.001	(-5.92,-4.31)
Special Education	-7.77*	.39	-2.11	<.001	(-8.53,-7.01)
English Learner	1.47*	.47	3.14	.002	(.55,2.39)
Free and Reduced Lunch	-2.22*	.30	-7.34	<.001	(-2.82,-1.63)
Mobile	-4.23*	.69	-6.17	<.001	(-5.58,-2.89)

*Indicates statistical significance at the p<.05 level

Source: Student Records and Vendor Data

Appendix B. Data Handling at UEPC

Data Collection Channel

The UEPC set up a dedicated secure file transfer protocol server for software vendors. All data exchanges between the UEPC and the vendors, schools, school districts, and USBE were compliant with FERPA and other federal and local privacy and confidentiality laws and regulations.

Data Disposition

All data that the UEPC received and derived from the received data will be used solely for this project and will be kept until the project ends. The UEPC will not share the linked data to any third party under any circumstances. The UEPC will not share any data components to any third party without formal written authorization by those who own the data components along with documentation of IRB approval from the third party's institution.

Once the project ends, all data will be sanitized and destroyed following the guideline of the University of Utah (http://regulations.utah.edu/it/guidelines/G4-004N1.pdf) and the Federal regulations (http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf, pp 22-23).

Data Source Security

All data were securely encrypted, transmitted, and stored according to industry and University of Utah standards.

Data Sources

Vendor Data

Seven math learning platforms were included in the evaluation, including ALEKS, Dreambox, Imagine Math, i-Ready, Mathscpace, MobyMax, and ST Math. Student usage from vendors was requested every month for analyses starting in September 2018 and going through May 2019.

USBE Data

In accordance with the existing data-sharing agreement, the USBE data needed for the evaluation of the software were transferred to the UEPC via the USBE's secure FTP server.

Data Storage

The Utah Education Policy Center (UEPC) considers the security and protection of data to be of the utmost importance. Encrypted data are stored on secure hardware, maintained by highly trained computer professionals, and safeguarded by the University of Utah's network security, Virtual Private Network (VPN), and firewall. The UEPC protects data in compliance with the Family Educational Rights and privacy Act, 20 U.S. Code §1232g and 34 CFR Part 99 ("FERPA"), the Government Records and Management Act U.C.A. §62G-2 ("GRAMA"), U.C.A. §53A-1-1401 et seq, 15 U.S. Code §§ 6501-6506 ("COPPA") and Utah Administrative Code R277-487 ("Student Data Protection Act").

The UEPC limits and restricts data access to leaders in charge of the day-to-day operations of the research, and professional and technically qualified staff who conduct research. All UEPC staff receive FERPA and CITI trainings and certification, which cover issues of data privacy, security, and protections, and ethics of data management and use. UEPC employees who have access to data are required to sign a Non-Disclosure Agreement. Access to data is controlled by password protection, encryption, dual authentication, and/or similar procedures designed to ensure that data cannot be accessed by unauthorized individuals.

The UEPC maintains a data sharing agreement (DSA) with the Utah State Board of Education (USBE) wherein the USBE shares data with the UEPC for the purposes of state, district, and federal evaluations.

Quoted from: Owens, R., Rorrer, A., Ni, Y., Onuma, F. J., Pecsok, M., & Moore, B. (2020). *Longitudinal Evaluation of the Math Personalized Learning Software Grant Program.* Salt Lake City, UT: Utah Education Policy Center.