National Math + Science Initiative's Rural ACCESS: AP, College, and Career Excellence in STEM and Computer Science Program Impact Study

Richard S. Brown, Ph.D.
West Coast Analytics

January 2025

Contents

Rural ACCESS: AP, College, and Career Excellence in STEM and Computer Science Program

Background

The National Math and Science Initiative (NMSI), a 501(c)(3) nonprofit organization, proposed an early-phase Education Innovation and Research (EIR) project, Rural ACCESS: AP, College, and Career Excellence in STEM and Computer Science, to increase the number of rural students engaged in and prepared for postsecondary Science, Technology, Engineering, and Math (STEM) coursework and accelerate their entry into STEM careers by deploying an innovative blended delivery model of NMSI's proven College Readiness Program (CRP).

NMSI was formed to address one of this nation's greatest economic threats — the declining number of students prepared to take rigorous college courses and pursue careers in STEM. CRP demonstrates that more students, especially high-need students, can master Advanced Placement STEM coursework by helping schools become centers of college readiness and career exploration. In recognition of the proven approach, CRP has received two Investing in Innovation Funding (i3) grants to help reach more than 1,000 schools and drive 175,000 AP qualifying scores (3 or higher on a 5-point scale). During this project, NMSI partnered with many of North Dakota's rural high schools (those located in rural locale codes) to grow AP enrollment and increase the number of qualifying scores at each school.

Ensuring all students have access to and excel in STEM is essential for our nation's economic growth (Langdon 2012). Growth of the U.S. STEM job market has far outpaced production of STEM degrees. This makes U.S. STEM dependent on foreign talent (US News and World Report 2016). Our STEM knowledge capital, which fuels innovation and ensures economic competitiveness, is at risk. Rural students are particularly underprepared for STEM. Students who succeed in rigorous, advanced academics in high school are more likely to pursue and complete four-year college degrees, placing them on a trajectory for high-potential economic and employment (College Board 2014). STEM, and particularly computer science, career fields are among the most lucrative in the U.S. but require postsecondary education (US News and World Report 2016). Yet many rural high schools can't offer any AP STEM due to acute teacher shortages; on average, 62 percent of U.S. rural schools experience a STEM teacher vacancy (Player 2015). This puts rural students at a significant disadvantage in developing college readiness skills, mindsets, and habits. In 2015, only 19 percent of the rural population held a bachelor's degree (US Department of Agriculture 2017), compared to the 33 percent urban average (Pew Research Center 2014).

NMSI's CRP increases the number and diversity of students taking and earning qualifying scores in AP STEM, by transforming partner schools into centers of college readiness. CRP makes a dramatic difference by supporting and motivating more youths to pursue postsecondary education with a focus on STEM and computer science in underserved rural regions.

This project builds on a proven CRP approach while addressing historical barriers to implementation within the rural context. Innovative strategies include 1) deploying a blended online delivery model, 2) leveraging existing relationships within rural communities to support access to and adoption of this model, 3) building the capacity of STEM teachers in rural areas to support CRP, and 4) reducing the program cost to participate in CRP.

Prior Research

A collection of research studies presents evidence of CRP's effectiveness, from impact on immediate outcomes related to AP to postsecondary results to longer-term, lifelong impacts. Holtzman (2010), using a comparative interrupted time series (CITS) design, found that in its first year, CRP had a positive and statistically significant impact on student enrollment in AP courses in mathematics, science, and English, as well as on their success in passing the related AP examinations with a score of 3 or higher. Notably, CRP implementation was associated with a 12-point increase in the percentage of students taking at least one mathematics, science, or English AP exam, showing growth of more than a full standard deviation.

Sherman (2014, 2015) provides longer-term evidence of CRP success, showing positive impacts on students' AP performance based on multiple years of program implementation across two cohorts of schools in Colorado and Indiana. Using a CITS design, they compared the changes in average AP outcomes over time of high schools implementing CRP (N=18) against changes in matched comparison schools that were not implementing the program (N=18). The study's first-year outcomes showed that CRP schools significantly outperformed the comparison schools both in the percentage of students taking AP STEM exams and in the percentage of students earning qualifying exam scores in these subjects. In the second year, treatment schools significantly outperformed comparison schools in the percentage of students taking AP exams and the percentage earning qualifying scores across all subject areas.

Jackson found that the program had positive effects on AP course enrollment, SAT/ACT scores, and college matriculation (Jackson 2007), as well as on college GPAs and college persistence (Jackson, 2010). Jackson (2014) also related CRP participation to enduring labor-market outcomes, such as wages. Brown & Choi (2015) employed a potential outcomes modeling approach on a large sample of treatment schools (N=287) to estimate the causal effect of CRP program participation on first-, second-, and third-year improvements over base year in AP exam-taking and AP qualifying scores. Their results indicate substantial and significant increases in both AP exam-taking and qualifying score-earning for all students, female students, and other student subgroups who have historically been underrepresented in STEM, when analyzed separately (average effect size: 0.64). More recently, Sherman (2017) deployed a CITS design to study the implementation of CRP across 58 high schools in Colorado and Indiana. Schools implementing CRP demonstrated significantly larger increases in the share of students taking and passing AP tests in targeted areas relative to comparison schools, and, importantly, gains were sustained over time.

Impact of blended modality: Online and blended learning programs are still relatively new to the education sector, and their design elements and application vary widely. In addition to the strong evidence that supports the CRP model, other research suggests that the method in the delivery method is irrelevant; it is the quality of the curriculum delivered and the effectiveness of the delivery method that affects performance (Clark 2012). Clark argues that the instructional method is important to generating learning, not the medium through which it is delivered, stating, "All methods required for learning can be delivered by a variety of media and media attributes" (Clark 2012; p. 181). Consistent with this perspective, we expect the positive effects of NMSI's CRP program will persist in an adapted delivery model.

The emergent literature on online and blended learning (OBL) programs suggests that effectiveness is tied to program design, content, and promising practices related to delivery (Amaka 2017). Promising practices related to delivery include interactivity, navigability, (a)synchronicity, flexibility, media richness, ease of use, individualization, mobility, proximity and responsiveness" (Amaka 2017). Blended CRP addresses these attributes. By incorporating proven content with evidence-based OBL formats, we expect that NMSI's blended delivery CRP will drive the same student outcomes as seen in standard delivery while greatly increasing the volume of high-need schools that NMSI can serve.

Impact Study

Study Description

Research Questions for the Study

The research questions driving this impact evaluation are:

- What is the effect of one year of school's implementation of CRP on the number of high school students enrolling in an AP STEM course compared to the proportion of students enrolling in an AP STEM course in the business-as-usual condition?
- What is the effect of one year of school's implementation of CRP on the number of high school students enrolling in an AP English course compared to the proportion of students enrolling in an AP English course in the business-as-usual condition?
- What is the effect of CRP on the number of high school students earning a qualifying score of 3 or better on an AP STEM test compared to students in the business-as-usual condition?
- What is the effect of one year of school's implementation of CRP on increasing the number of high school students enrolling in an AP STEM course compared to the increase in the number of students enrolling in an AP STEM course in the business-as-usual condition?
- What is the effect of one year of school's implementation of CRP on increasing the number of high school students enrolling in an AP English course compared to the increase in the number of students enrolling in an AP English course in the business-as-usual condition?

- What is the effect of CRP on increasing the number of high school students earning a qualifying score of 3 or better on an AP STEM test compared to the increase in the number of students in the business-as-usual condition?

Intervention Condition

The Rural Access program, implemented by the National Math and Science Initiative (NMSI), sought to address significant educational barriers in rural U.S. schools by increasing access to Advanced Placement (AP) STEM courses, enhancing college readiness, and fostering economic opportunities for underserved students. Over four years, the initiative provided a blended learning model that combined remote AP courses, teacher training, and tailored student supports (see Figure 1 for Logic Model).

In its inaugural year (2018-2019), the program focused on planning and piloting the blended College Readiness Program (CRP) in 12 rural North Dakota schools. The initial phase involved developing interactive online modules for AP courses, training AP and pre-AP teachers, and establishing foundational evaluation metrics. NMSI partnered with Virtual High School (VHS) Learning to provide AP courses that were otherwise unavailable in these communities. Despite challenges in adapting students to independent online learning and low engagement with coaching resources, the pilot successfully enabled schools to offer AP Computer Science Principles for the first time and served 841 students. However, retention rates were lower than anticipated. Observations from focus groups highlighted the need for more robust student support systems to navigate the new online learning model.

Figure 1. Logic Model for Blended CRP

	Inputs	Outputs	Outcomes	Impacts
Support	• Course Specific Training (Content Workshop/ Summer Institute) • NMSI year-round supports (PD offerings, mentorship, online curricular resources)	• AP teachers trained at partner schools • Pre-AP teachers trained in feeder pattern for partner schools	• AP teachers report knowledge and implementation of high-quality content/ instruction in their classroom	• Prepared and aligned teacher team at partner schools
Support Students	• Serve high needs students • Student supports (study sessions, content reviews, student coaches)	• Increase from baseline year AP enrollment	• Greater # and % of students taking AP exams • Greater # and % of students receiving qualifying scores on exams	• Pipeline of AP-ready students • Improved college readiness and matriculation
Support Schools	• Serve rural schools	• AP courses prioritized and sequenced in partner schools • Barriers to AP eliminated at rural partner schools	• Greater # and % of schools offer AP • Greater AP enrollment at schools • Increase in # and % of qualifying scores from baseline at schools	• Culture of improvement/ achievement among leaders, teachers, students, and parents • STEM Learning is valued/promoted among leaders, teachers, students, and parents

Year 2 (2019-2020) marked the program's first full implementation, with a focus on expanding access and refining the delivery model. NMSI partnered again with VHS Learning to expand AP course offerings, enabling schools like Langdon Area High School to introduce on-campus AP Biology after teacher training supported by the program. Professional development efforts emphasized building local capacity, equipping teachers with skills to independently sustain AP coursework. Unfortunately, the COVID-19 pandemic disrupted the latter half of the year, presenting significant challenges to both recruitment and retention. The number of participating schools dropped from 12 to 9, and the total number of students served fell to 547. However, the program achieved cost efficiencies. Students provided overwhelmingly positive feedback about the integrated online platform, but pandemic-induced stress and lack of in-person support contributed to higher dropout rates, especially among online learners.

During Year 3 (2020-2021), the program navigated the prolonged effects of the pandemic while expanding its reach to schools in New Mexico, Pennsylvania, and Vermont. Key focus areas included improving retention and adapting to new challenges in virtual education. Retention improved significantly overall, with 85% of students persisting, though online-only learners continued to struggle, with a retention rate of 43.5%. The program also succeeded in supporting partner schools to establish new on-campus AP courses, reducing reliance on external providers.

For instance, schools were able to offer AP Chemistry and AP English Literature on-site for the first time. Despite exceeding AP enrollment targets by 50%, exam performance did not meet expectations, reflecting a national trend of lower scores during the pandemic. Innovations such as granting school administrators direct access to gradebooks facilitated better monitoring and student support. However, wellness checks during coaching sessions revealed heightened mental health challenges among students, underlining the broader impacts of the pandemic.

In its final year (2021-2022), the program prioritized scaling efforts and ensuring sustainability. While it fell short of its goal to engage 20 schools, with 19 actively participating, it achieved strong on-campus student retention rates exceeding 90%. Teacher training initiatives expanded significantly, with the introduction of NMSI's Laying the Foundation (LTF) program, targeting both middle and high school educators. This effort aimed to strengthen the pipeline of AP-ready students and prepare teachers to lead advanced coursework in the future. Persistent challenges included a 55% withdrawal rate for online learners and reduced engagement due to "Zoom fatigue." Despite these hurdles, the program made strides in influencing broader educational initiatives, including Alabama's Rural Learning Accelerator and a pilot program for military-connected students. These spin-offs highlighted the transferability of the Rural Access model to other underserved populations. Additionally, NMSI began exploring opportunities to support Native American communities in New Mexico, identifying shared challenges in teacher recruitment and student access to advanced courses.

Overall, the Rural Access program demonstrated the critical role of on-site support in ensuring student success, particularly in rural areas where personal relationships with educators are pivotal. The initiative highlighted the importance of building local teacher capacity for sustainable AP offerings and underscored the challenges of engaging students in virtual learning environments. Despite the obstacles encountered, the program made substantial progress in addressing educational inequities, enabling rural students to access rigorous coursework and fostering college readiness. The lessons learned have informed broader educational policies and provided a model for innovative approaches to STEM and AP education, laying the groundwork for future efforts to support underserved communities.

Setting

This study took place across a several states including North Dakota, Pennsylvania, New Mexico, Texas, and Vermont. All of the schools participating in the study are in rural settings.

Comparison Condition

The comparison schools were identified and selected based on a propensity score matching (PSM) approach. The approach employed in this study is the Nearest Neighbor Matching approach. Nearest Neighbor Matching (NNM) is a commonly used technique within propensity score matching, a statistical approach for estimating causal effects in observational studies. The primary goal of PSM is to mimic the conditions of a randomized controlled trial by balancing observed covariates between treated and untreated groups. This balance is achieved by matching units with similar propensity scores, which represent the probability of receiving treatment given a set of observed covariates. In NNM, each treated unit is matched to the untreated unit with the closest propensity score, ensuring that the groups are comparable in terms of their observed characteristics.

The process of NNM begins with estimating propensity scores using methods like logistic regression or machine learning models. Once scores are calculated, matching is performed based on proximity in propensity scores, with options for matching with or without replacement. Matching with replacement allows a control unit to be matched to multiple treated units, potentially improving match quality but increasing variance. In contrast, matching without replacement preserves the sample size but can lead to suboptimal matches. To improve the quality of matches, a caliper—a threshold for the maximum allowable difference in propensity scores—can be applied to exclude poorly matched pairs.

After matching, the balance of covariates between treated and untreated groups is assessed to ensure comparability, often using metrics like standardized mean differences or visual tools such as density plots. Once balance is confirmed, treatment effects are estimated by comparing outcomes between the matched groups, providing insights into the causal impact of the intervention. NNM can be extended to include one-to-many or many-to-many matching, where multiple units are matched, often with weights assigned based on the closeness of propensity scores.

The advantages of NNM lie in its simplicity, interpretability, and ability to improve covariate balance by focusing on a single balancing score. However, it has limitations, including potential sample size loss due to unmatched units and sensitivity to the specification of the propensity score model. Poorly estimated propensity scores or the absence of caliper thresholds can lead to biased estimates if poorly aligned matches are included. Additionally, the method involves a tradeoff between bias and variance, particularly when choosing between matching with or without replacement.

NNM is widely used in various fields such as healthcare, economics, and social sciences to evaluate the effects of treatments, policies, and programs. By addressing confounding factors in observational data, NNM enhances the credibility of causal inferences and remains a vital tool for researchers aiming to draw robust conclusions from non-experimental data. Despite its simplicity, careful implementation and validation of assumptions are critical to ensure reliable results.

For this study, data from the National Center for Educational Statistics (NCES) and The College Board for all public schools in Texas, Pennsylvania, and Georgia were used to identify schools similar to the treatment schools in characteristics such as rural code, enrollment totals, percentage of students that are Black or Hispanic, and percentage of student who qualified for Free or Reduced Price Lunch (FRPL). These characteristics have been shown to correlate with the AP outcomes measures of interest in this study.

Using the NNM approach described above, a comparison school most similar to each of the 27 treatment schools was identified and assigned as a comparison school for analysis. These schools were also reviewed to ensure that they had not participated in the NMSI CRP program at any point in the past. The comparison group is considered a “business as usual” condition.

Study Participants

Treatment schools were identified and recruited to participate by NMSI personnel.

During this first implementation period, NMSI completed its first full academic year of program implementation of "Rural Access: AP, College, and Career Excellence in STEM and Computer Science"—an initiative designed to ensure that students in rural schools, which often lack the resources and critical mass of students and teachers necessary to maintain a robust Advanced Placement framework, have access to the advanced STEM coursework that research shows fosters college readiness and in turn advances students' economic opportunities.

This program mirrors NMSI's College Readiness Program (CRP), which offers teacher professional development and direct-to-student supports but for the purposes of this grant was modified to test a blended learning environment and include remote access to entire AP courses, offered by a partner of NMSI, VHS Learning. They began the 2019-2020 year with 12 treatment schools, though three dropped out of the program when their Advanced Placement students were unable to continue in AP courses. Nine program schools completed the 2019-20 school year as part of this grant.

For the purposes of this grant, NMSI defines high-need students as those enrolled in rural schools with an NCES locality code 41, 42, or 43. The premise is that rural students face significant barriers in accessing rigorous STEM courses, regardless of socioeconomic status. Acute teacher shortages put rural students at a disadvantage in developing college readiness, a need this grant seeks to ameliorate by introducing a range of online AP courses for schools that are unable to offer them onsite and creating online coaching spaces and supports for rural students. In addition to the online courses, which increase access to advanced STEM coursework, this grant supports teacher training to develop the capacity in program schools to offer in- person AP courses.

The original project goal was to implement the program in 20 schools with NCES rural locality codes 41, 42, and 43. Over the course of the grant, NMSI recruited 31 schools in 10 states, but retention was a consistent challenge. The long-term effects of COVID-19 continued to present significant challenges to students, reducing our ability to retain schools over the course of the grant. The small size of many potential partner schools also made the overall goal of implementing the program in a certain number of schools difficult to meet and measure—often partner schools had only one student who engages with the program, and if that student chooses to drop his or her AP class then the school is no longer an active participant for the year. In such cases the schools were excluded from the impact study, along with two private schools that could not be matched on the propensity matching. During this grant period 27 public schools were engaged and 16 provided data for the impact analysis. The implementation schools are listed in the Table below. The shaded rows indicate schools that provided outcome data for the impact analyses.

Table 1. Treatment Schools

STATE	School	Enrollment
NEW MEXICO	EAST MTN HIGH SCHOOL	366
NEW MEXICO	CLOUDCROFT HIGH	126
NEW MEXICO	DORA HIGH	113
NEW MEXICO	FORT SUMNER HIGH	97
NEW MEXICO	MELROSE HIGH	55
NORTH DAKOTA	LINTON HIGH SCHOOL	90
NORTH DAKOTA	RICHLAND JUNIOR-SENIOR HIGH SCHOOL	139
NORTH DAKOTA	LITCHVILLE-MARION HIGH SCHOOL	54
NORTH DAKOTA	ENDERLIN AREA HIGH SCHOOL	159
NORTH DAKOTA	ELLENDALE HIGH SCHOOL	141
NORTH DAKOTA	HEBRON HIGH SCHOOL	74
NORTH DAKOTA	KULM HIGH SCHOOL	48
NORTH DAKOTA	LANGDON AREA HIGH SCHOOL	186
NORTH DAKOTA	LARIMORE HIGH SCHOOL	210
NORTH DAKOTA	WATFORD CITY HIGH SCHOOL	508
NORTH DAKOTA	WILTON HIGH SCHOOL	104
NORTH DAKOTA	MONTPELIER HIGH SCHOOL	54
NORTH DAKOTA	RAY HIGH SCHOOL	81
NORTH DAKOTA	ROLETTE HIGH SCHOOL	50
PENNSYLVANIA	BENTWORTH SHS	358
VERMONT	U-32 HIGH SCHOOL (UHSD #32)	775
TEXAS	SILVER VALLEY HIGH	397
TEXAS	CRAWFORD COUNTY HIGH SCHOOL	462
TEXAS	TURTLE MOUNTAIN COMMUNITY HIGH SCHOOL	510
TEXAS	CLAYTON HIGH	124
TEXAS	MORENO VALLEY HIGH	55
TEXAS	ONATE HIGH	1497

Table 2. Comparison Schools

STATE	School	Enrollment
PENNSYLVANIA	BELLE-VERNON AREA HS	843
PENNSYLVANIA	BELLWOOD-ANTIS HS	392
PENNSYLVANIA	BLACKLICK VALLEY JSHS	265
PENNSYLVANIA	DANVILLE AREA SHS	606
PENNSYLVANIA	FAIRFIELD AREA HS	369
PENNSYLVANIA	MERCER AREA SHS	410
PENNSYLVANIA	NEWPORT HS	311
PENNSYLVANIA	SALISBURY-ELK LICK JSHS	143
PENNSYLVANIA	SOUTH SIDE HS	395

PENNSYLVANIA	SOUTHERN COLUMBIA HS	408
PENNSYLVANIA	UNION HS	257
PENNSYLVANIA	UPPER DAUPHIN AREA HS	386
TEXAS	MIDWAY SCHOOL	97
TEXAS	MILANO H S	156
TEXAS	NEW DIANA H S	290
TEXAS	ORANGEFIELD H S	530
TEXAS	PORT ARANSAS H S	161
TEXAS	S AND S CONS H S	268
TEXAS	JAMES BOWIE H S	165
TEXAS	SMYER H S	183
TEXAS	WAXAHACHIE GLOBAL H S	418
TEXAS	GOLIAD H S	407
TEXAS	HICO H S	279
TEXAS	WALL H S	334
TEXAS	WELLINGTON H S	151
TEXAS	LAMAR CONS H S	1655
TEXAS	ONALASKA JR/SR HIGH	399

Design and Measures

Independence of the Impact Evaluation

This Impact Evaluation is being conducted by West Coast Analytics, independent from the National Math + Science Initiative. The data for the comparison samples was provided directly to West Coast Analytics from The College Board. The data for the treatment schools was provided by the participating schools. All data regarding the treatment schools and participation was provided to West Coast Analytics by NMSI.

Pre-registration of the Study Design

Initially, the study design involved a multi- quasi-experimental (QE) study using a Comparative Interrupted Time Series (CITS). Comparison schools will be selected using propensity score-matching techniques. During each year of study implementation (2019-2020, 2020-2021), the CRP impact was going to be evaluated using CITS design (CITS; Shadish, Cook, & Campbell, 2002). In this design, the plan was to examine the change in the program schools' performance using student-level outcomes, when the program was implemented, benchmarked against the change for a similar set of comparison schools. Up to 30 North Dakota and Louisiana high schools were to be recruited for the study, and an equal number of schools were to be identified as comparison schools matched using propensity score matching techniques. All schools receiving the treatment were to be used in the estimate of program effects. This plan was pre-registered. However, due to problems with recruitment and retention this plan had to be amended.

Design

This study still uses a quasi-experimental design (QED), but not a CITS design. Rather, this study combines all schools into a single analysis contrasting the baseline year outcomes with the outcomes following the first year of implementation since not all of the schools implemented in the same year and few implemented in the initial years of the study. The analysis involves a group comparison of outcomes between the treatment and comparison school groups on the treatment outcomes after baseline equivalence is established.

Measures

The outcome measures used in this study are counts of Advanced Placement test scores in English; counts of Advanced Placement test scores in STEM courses (Math and Sciences combined), and counts of Advanced Placement test scores in STEM courses having a value of 3 or greater, indicating was is considered a Qualifying Score on the AP Exam. For each of these outcome measures, the AP examinations are administered and scored by The College Board. The AP examinations are administered in the Spring at the same time for all schools, either treatment or comparison.

Data Analysis and Findings

Baseline Equivalence

To estimate baseline equivalence, we compared the treatment and comparison groups on the variables used for the propensity matching, which included total enrollment, percentage of students that were Black or Hispanic, and percentage of students that qualified for Free or Reduced-Price Lunch (FRPL). We computed treatment-control differences in the mean values for each of the groups on each of the matching variables and conducted analysis of variance for each comparison. The treatment and comparison groups did not statistically differ on total enrollment ($F(1,53) = 2.35$; $p = .131$); percentage of students that were Black or Hispanic ($F(1,53) = .004$); $p = .951$); or percentage of students who qualified for FRPL ($F(1,53) = .003$; $p = .956$).

Table 3. Baseline Equivalence Assessment – Matching Variables Full Sample

	Comparison Group			Treatment Group				
Measure	Sample Size	Mean	Standard Deviation	Sample Size	Mean	Standard Deviation	Treatment – Control Difference	Standardized Difference
Total Enrollment	27	380.67	301.04	27	253.07	310.58	-127.6	0.41
% Black or Hispanic	27	15.64	19.88	27	15.98	20.31	.34	0.02
% FRPL	27	32.45	11.14	27	32.27	11.54	-.18	-0.02

We then conducted a similar baseline equivalence analysis just for the sample providing outcome data for the impact analysis. Similarly, we found no statistical differences on the matching variables. We computed treatment-control differences in the mean values for each of the groups on each of the matching variables and conducted analysis of variance for each comparison. The treatment and comparison groups did not statistically differ on total enrollment ($F(1,31) = 0.936$; $p = .341$); percentage of students that were Black or Hispanic ($F(1,31) = .001$); $p = .979$); or percentage of students who qualified for FRPL ($F(1,31) = .002$; $p = .965$).

Table 4. Baseline Equivalence Assessment – Matching Variables Analytic Sample

	Comparison Group			Treatment Group				
Measure	Sample Size	Mean	Standard Deviation	Sample Size	Mean	Standard Deviation	Treatment – Control Difference	Standardized Difference
Total Enrollment	16	433.06	351.71	16	316.06	342.88	-117.0	-0.33
% Black or Hispanic	16	19.25	22.61	16	19.03	22.90	-0.22	-0.01
% FRPL	16	28.54	10.97	16	28.71	11.62	0.17	0.02

Having established no significant difference on the matching variables, we then compared the two groups on the outcome measures in the year preceding implementation of the CRP program (baseline year). Again, we provide sample descriptives in terms of means and standard deviations, and computed treatment-comparison differences and standardized differences These are provided in the table below. We also conducted analysis of variance to test the statistical significance of the mean differences. We found no significant differences in the baseline outcome measures. For AP English exams the treatment-control difference in means was 11.94 but was not significant ($F(1,31) = .979$; $p = .330$). The treatment-control difference for AP STEM exams was 15.31 but also not significant ($F(1,31) = .517$; $p = .478$). Finally, the treatment-control difference in AP STEM Qualifying scores was -4.31 and non-significant ($F(1,31) = .193$; $p = .664$).

Table 5. Baseline Equivalence Assessment – Outcomes at Baseline Year

	Comparison Group			Treatment Group				
Measure	Sample Size	Mean	Standard Deviation	Sample Size	Mean	Standard Deviation	Treatment – Control Difference	Standardized Difference
AP English	16	14.19	41.43	16	26.13	41.43	11.94	0.35
AP STEM	16	27.13	49.56	16	42.44	69.33	15.31	0.26
AP STEM Qualifying Scores	16	17.56	34.24	16	13.25	19.24	-4.31	-0.16

Program Effects

To estimate program effects, we conducted analysis of variance on the outcomes of interest comparing the mean values of the outcome measures for the treatment and control groups. In this study, there was considerable attrition in the treatment group, as only 16 schools assigned to the treatment group out of 27 originally implementing the program provided outcome data for analysis. Data were not imputed for these schools, rather they were excluded from the analysis. Subsequently these results may not be indicative of the impact had these schools provided outcome data. The matched comparison schools for those treatment schools were retained in the analysis but the matched comparison schools for the treatment schools not providing outcome data were removed from the analytic sample. Thus, the program impact analysis is restricted to a treatment group comprised of 16 schools and a comparison group comprised of 16 matched comparison schools to those treatment schools.

We looked at two types of treatment effects. The first is the effect of participating in the program on Advanced Placement test counts in English, STEM, and Qualifying STEM scores. The second is the effect of participating in the program on gains (from baseline to year after implementation) in AP Test counts in English, STEM, and AP STEM Qualifying scores.

For the first analysis we find considerable variation in the performance across both treatment and comparison schools, as indicated by high standard deviations for each outcome measure. For example, the mean number of AP English exams in the treatment condition for the year following implementation was 42.13, but the standard deviation was 69.23. Similarly, the mean number of AP STEM exams in the post-implementation period was 87.31 for the treatment group, but the standard deviation for these 16 schools was 112.94. As a result of the large within group variation for this small sample of treatment and comparison schools, between group differences of statistical significance are hard to identify. However, we do observe a few interesting findings. For example, the post-implementation treatment-control difference for AP STEM test counts exceeds 58. This indicates that on average treatment schools had 58 more AP STEM students than their comparison schools. That difference was only about 15 at baseline. And while this post-implementation difference in means does not reach traditional levels of statistical significance ($F(1,31) = 3.46$; $p = .073$) at a 95% confidence level, it does do so at a 90% confidence level. Given the small number of participating schools and the substantial within group variation, this finding is suggestive of a positive impact of the CRP program on AP STEM participation.

Table 6. Impact Analysis Results for AP Test Counts

Outcome Measure	Comparison Group				Treatment Group				Treatment – Control Difference	Standardized Difference	p-value
	Sample Size				Sample Size						
	# Clusters	# Students	Mean	Standard Deviation	# Clusters	# Students	Mean	Standard Deviation			
AP Eng Exams	16	5373	13.43	23.87	16	6929	42.13	69.23	28.70	0.54	.128
AP STEM Exams	16	5373	28.88	55.25	16	6929	87.31	112.94	58.43	0.63	.073*
AP STEM Qualifying Scores	16	5373	18.75	40.46	16	6929	34.56	47.91	15.81	0.36	.321

The more compelling findings deal with the impact on gains in the outcome measures. The whole premise of the NMSI program is that participation in the program leads to more opportunities and participation for the underserved communities in STEM courses and opportunities. Thus, it is expected that schools participating in the program would gain in the areas of STEM participation and STEM preparedness. This is exactly what the data suggest. For example, participating schools increased the number of AP English exams more than comparison schools by an average more than 16 exams ($F(1,31) = 3.63$; $p = .067$). Again, though not statistically significant at the 95% confidence level, it is at the 90% confidence level, suggesting a positive impact. Similarly, the gain in qualifying STEM scores for treatment schools was on average more than 20 per school more than comparable schools ($F(1,31) = 3.31$; $p = .079$).

The strongest, and most significant effect of the CRP program is on the outcome the program is most intended to influence. Treatment schools significantly improved the number of students taking AP STEM courses relative to matched comparison schools. On average treatment schools increased AP STEM participation by an average of more than 44 students per school compared to only 2 per school in the comparison condition. This was the most statistically significant impact finding in this study ($F(1,31) = 5.04$; $p = .032$), and is consistent with the intent of the program.

Table 7. Impact Analysis Results for AP Test Gains (After Implementation-Baseline)

Outcome Measure	Comparison Group				Treatment Group				Treatment – Control Difference	Standardized Difference	p-value
	Sample Size		Mean	Standard Deviation	Sample Size		Mean	Standard Deviation			
	# Clusters	# Students			# Clusters	# Students					
AP Eng Exam Gains	16	5373	-.50	4.66	16	6929	16.0	34.35	16.5	0.65	.067*
AP STEM Exam Gains	16	5373	2.06	17.31	16	6929	44.88	74.26	42.82	0.75	.032**
AP STEM Qualifying Score Gains	16	5373	1.19	15.76	16	6929	21.31	41.37	20.12	0.62	.079*

Discussion

The evaluation of the National Math + Science Initiative's (NMSI) Rural ACCESS: AP, College, and Career Excellence in STEM and Computer Science Program highlights the potential of innovative, blended learning models to address long-standing inequities in STEM education among rural high school students. The program aimed to provide access to Advanced Placement (AP) STEM courses, enhance college readiness, and foster career opportunities for underserved students in rural communities, where access to advanced coursework and qualified teachers is often limited.

This study's quasi-experimental design established baseline equivalence between treatment and comparison schools on key variables, ensuring credible comparisons. The impact analysis revealed several key findings. Statistically significant gains in AP STEM exam participation ($p = 0.032$) and trends toward significant gains in AP English exams and qualifying STEM scores, suggest that the program effectively increased access to and success in AP coursework. Further, substantial variability within groups, reflecting diverse outcomes across schools, likely influenced by contextual factors such as school size, teacher availability, and community engagement, made it more challenging to find statistical effects. Nonetheless, the program demonstrated a substantial positive effect on increasing the number of AP STEM exams taken by students in treatment schools compared to comparison schools. The mean difference in AP STEM participation gains (44.88 students per treatment school compared to 2.06 students in comparison schools) underscores the program's effectiveness in expanding access to STEM education.

The program also demonstrated an ability to deliver professional development in rural contexts. A cornerstone of the initiative was the professional development of rural educators, enabling them to sustain AP coursework independently. While teacher training was effective in building local capacity, schools' reliance on external providers for certain AP courses highlighted the ongoing challenges of rural teacher shortages. The blended approach, combining online and in-person learning, proved effective in expanding access to AP courses. However, pandemic-induced disruptions and challenges with online learning (e.g., "Zoom fatigue") limited its full potential. Feedback from students highlighted the importance of interactivity and robust support structures in online learning environments.

However, retention rates for online learners remained a persistent issue throughout the program. The 55% withdrawal rate for online courses, coupled with the 43.5% retention rate in the third year for virtual learners, emphasizes the need for more robust student support systems and personalized engagement strategies. Much of this can be attributed to the disruption caused by the pandemic just as this project began. The COVID-19 pandemic significantly disrupted program implementation, affecting student participation, retention, and AP exam performance. Despite these challenges, the program's flexibility allowed it to adapt and continue serving rural communities.

The Rural ACCESS program made notable strides in addressing systemic barriers to STEM education in rural areas. By leveraging a blended delivery model and prioritizing teacher training, the initiative expanded AP STEM access and participation, setting rural students on pathways toward higher education and lucrative STEM careers. Key accomplishments include: increased AP STEM exam participation; enhanced teacher capacity to deliver AP coursework, fostering long-term sustainability; and broadened access to advanced coursework in underserved rural communities, despite challenges such as high withdrawal rates for online learners and pandemic-related disruptions.

However, the study underscores the need for strengthened student support systems, particularly for online learners, to address engagement and retention challenges. There is also a need for continued investment in rural teacher training and recruitment to build sustainable local capacity. Lastly, this study shows that adaptation of program strategies to mitigate external challenges such as those posed by the COVID-19 pandemic are possible. Ultimately, the lessons from this evaluation provide valuable insights for scaling and refining STEM initiatives in rural and underserved contexts, ensuring equitable access to high-quality education for all students.

Fidelity of Implementation Study

Fidelity Measurement

As indicated in Figure 1 presented previously, the NMSI program is comprised of three components: 1) teacher supports; 2) student supports, and 3) school supports. More specifically, with respect to the first component of teacher supports, the program is implemented by engaging teachers in the Summer Institute content workshops, Laying the Foundation training for pre-AP teachers, year-round mentorship and online supports, and increased teacher knowledge and use of content and instructional strategies. Implementation of student supports is determined by student use of online study sessions and study coaches. Implementation of school supports is measured by schools adding AP courses and/or altering course sequencing to facilitate expanded access and to prioritize student success as well as by increasing the use of online programs to facilitate expanded student access to AP courses in the school.

Table 8 shows the implementation metrics and thresholds for each component. For the first key component, teacher supports, implementation is measured by four elements. The first three elements are percentage of participating AP teachers attending the Summer Institute trainings, percentage of pre-AP teachers attending Laying the Foundation training, and the percentage of participating teachers utilizing the online supports. Each of these three elements has a target threshold of 80%. If 80% or more of the relevant participants attend the training or use the resources, this implementation target is met for that element. The fourth element in this key component is that participating AP teachers report an increase in their knowledge and usage of content and effective instructional strategies. For this element, the target threshold is 66%. To fully implement the program on this key component, the sum of these four element scores should be 3 or above.

For the second key component, student supports, there is only one implementation metric. This component is measured by the percentage of students in the participating teachers' courses that attend online study sessions and utilize student coaches. The threshold for this metric is 50%.

The school supports key component has two implementation measures. The first is the percentage of schools that add AP courses and/or alter courses sequencing to facilitate expanded access and prioritize student access to AP courses. The second is implementation measure for this key component is increased use of online programs to facilitate expanded student access to AP courses at the school. Both of theses implementation metrics have a threshold of 100%.

Table 8. Implementation of Each Key Component in Program Logic Model

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring Goal	Indicator Scoring at Sample Level
Key Component 1. Teachers Supports				
(1) AP teachers attend the College Readiness Program Summer Institute	Program	Percentage of participating teachers attending training	80%	1 = Meets goal 0 = Does not meet goal
(2) Pre-AP teachers attend Laying the Foundation training	Program	Percentage of participating teachers attend training	80%	1 = Meets goal 0 = Does not meet goal
(3) Participating AP teachers will take advantage of NMSI's year-round supports	Program	Percentage of participating teachers utilizing supports	80%	1 = Meets goal 0 = Does not meet goal
(4) Participating AP teachers report an increase in knowledge and use of content and effective instructional strategies	Program	Percentage of participating teachers reporting increase in knowledge and use of content and effective instructional strategies	66%	1 = Meets goal 0 = Does not meet goal
Key Component 1 Total Score **Professional Development**			Sum of school-level indicator scores (range=0-4) **Adequate score = 3**	Sum of sample-level indicator scores (Range = 0-4) **Adequate = 3**
Key Component 2. Student Supports				
(1) Students attend online study sessions and utilize student coaches	Teacher	Percentage of students attending online study sessions and utilizing student coaches	50%	1 = Meets goal 0 = Does not meet goal

Table 8. Implementation of Each Key Component in Program Logic Model

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring Goal	Indicator Scoring at Sample Level
Key Component 2 Total Score **Student Supports**			Sum of school-level indicator scores (range=0-1) **Adequate score = 1**	Sum of sample-level indicator scores (Range = 0-1) **Adequate = 1**
Key Component 3. School Supports				
(1) Participating schools add AP courses and/or alter course sequencing to facilitate expanded access and prioritize student success	Program	Percentage of participating schools that alter course sequencing and expand access to AP courses	100%	1 = Meets goal 0 = Does not meet goal
(2) Participating schools increase use of online programs to facilitate expanded student access to AP in a rural setting		Percentage of participating schools that increase use of online programs to expand AP access	100%	1 = Meets goal 0 = Does not meet goal
Key Component 3 Total Score **School Supports**		Sum of scores	Sum of school-level indicator scores (range=0-2) **Adequate school score = 2**	Sum of sample-level indicator scores (Range = 0-2) **Adequate = 2**

Fidelity Findings

For the project, the fidelity of implementation findings are mixed, with more difficulty in implementation in the early years, but improved implementation in the last two years. In the first year, only one of the three key components were implemented at the level of expectation. The teacher supports component had a score of 2 where 3 is the target. Although 100% of AP teachers attended Summer Institute, only 50% of pre-AP teachers attended Laying the Foundation training. 82% of teachers reported increased knowledge and use of content and instructional strategies, but not data were provided on use of online supports. The student supports target was met, with 59.5% of students utilizing supports. The school supports target was not met (Score of 0 with target score of 2) although 92% of schools added AP courses and increased use of online programs to facilitate student access to AP and 75% of schools increased use of online programs to facilitate student access to AP. The second year was the most Covid impacted year and no data on fidelity of implementation was provided.

The third year of implementation saw an incremental improvement in fidelity of implementation with two of the three key components fully implemented. Both student supports and school supports were fully implemented. The teacher supports component fell short of full implementation with a score of 2, although 86% of AP teachers attended Summer Institute and 92.5% indicated increased knowledge and use of content. However, only 42% used year-round supports and No pre-AP teachers were included or reported. The fourth and final year showed complete implementation fidelity, with all three key components hitting their targets. The teacher supports component yielded a score of 3 with 83% of teachers attending the Summer Institute and 96% indicating increased knowledge and use of content. Additionally, 75% used year-round supports. No pre-AP teachers were included or reported on. Student supports showed 90% of students attended online study sessions and utilized student coaches. For the school supports component, 100% of schools added AP courses and increased use of online programs to facilitate student access to AP courses in their rural settings.

In summary, the level to which the program was implemented with fidelity increased gradually from somewhat to complete implementation over the project period. As has been noted elsewhere in this report, the project and the implementation of the program was severely hampered by the onset of the Covid pandemic. Although initial implementation was wanting, gradual development in the third and fourth years to full implementation demonstrated considerable improvement in dealing with the Covid challenges and struggles associated with these rural environments.

Table 9. Findings on Fidelity of Implementation by Component by Year 1 and Year 2

Key Components, Number of Indicators, Units, and Threshold				Year 1 Results (2019-20 School Year)			Year 2 Results (2020-21 School Year)		
Key Component	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold for Fidelity of Implementation	Number of Units in Which Component was Implemented	Number of Units in Which Fidelity of Component was Measured	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold	Number of Units in Which Component was Implemented	Number of Units in Which Fidelity of Component was Measured	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold
1.Professional Development	**4**	**Program**	**Sample-level component score of at least 3 each year**	12 Schools	12 schools	Score is 2 ***Program fidelity = No***	**9 schools**	**Not reported**	Score is 0 ***Program fidelity = No***
2.Student Supports	1	Program	**Sample level score of 1 each year**	440 students 12 schools	262 students 12 schools	Score is 1 59.5% of students utilized supports ***Program fidelity = Yes***	9 schools	Not reported	Score is 0 ***Program fidelity = No***
3.School Supports	2	Program	**Sample level score of at least 2 each year**	12 schools	**11 schools**	Score is 0 92% of schools added AP courses and increased use of online programs to facilitate student access to AP ***Program fidelity = No***	**9 schools**	**Not reported**	Score is 0 81% of schools added AP courses and increased use of online programs to facilitate student access to AP ***Program fidelity = No***

Table 10. Findings on Fidelity of Implementation by Component by Year 3 and Year 4

Key Components, Number of Indicators, Units, and Threshold				Year 3 Results (2021-22 School Year)			Year 4 Results (2022-23 School Year)		
Key Component	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold for Fidelity of Implementation	Number of Units in Which Component was Implemented	Number of Units in Which Fidelity of Component was Measured	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold	Number of Units in Which Component was Implemented	Number of Units in Which Fidelity of Component was Measured	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold
1.Professional Development	**4**	**Program**	**Sample-level component score of at least 3 each year**	23 Schools **40 teachers**	23 schools **40 teachers**	Score is 2 86% of teachers attended Summer Institute and 92.5% indicated increased knowledge and use of content; only 42% used year-round supports and No pre-AP teachers were included ***Program fidelity = No***	**19 schools**	**19 schools**	Score is 3 83% of teachers attended Summer Institute and 96% indicated increased knowledge and use of content; and 75% used year-round supports. No pre-AP teachers were included ***Program fidelity = Yes***
2.Student Supports	1	Program	**Sample level score of 1 each year**	44 students	42 students	Score is 1 95% of students attended online study sessions and utilized student coaches ***Program fidelity = Yes***	32 students	29 students	Score is 1 90% of students attended online study sessions and utilized student coaches ***Program fidelity = Yes***
3.School Supports	2	Program	**Sample level score of at least 2 each year**	17 schools	**17 schools**	Score is 2 100% of schools added AP courses and increased use of online programs to facilitate student access to AP ***Program fidelity = Yes***	**19 schools**	**19 schools**	Score is 2 100% of schools added AP courses and increased use of online programs to facilitate student access to AP ***Program fidelity = Yes***

References

Brown, R., Choi, K. 2015. Measuring the Casual Effect of the National Math and Science Initiative's College Readiness Program. National Center for Research on Evaluation, Standards, and Student Testing.

Bureau of Labor Statistics. 2017. "69.7 percent of 2016 high school graduates enrolled in college in October 2016." TED: The Economics Daily. May 22. https://www.bls.gov/opub/ted/2017/69-point-7-percent-of-2016-high-school-graduates- enrolled-in-college-in-october-2016.htm.

Carnevale, A. and Rose, S,. 2015. The Economy goes to College: The Hidden Promise of Higher Education in the Post-Industrial Service Economy. Georgetown Center on Education and the Workforce.

Clark, R. 2012. Learning from Media: Arguments, Anlysis, and Evidence, Second Edition. University of Southern California.

College Board. 2016. Qualifying Scores Data.

College Board. 2014. 10 Years of Advanced Placement Exam Data Show Significant Gains in Access and Success; Areas for Improvement. https://www.collegeboard.org/releases/2014/class-2013-advanced-placement-results- announced.

College Board. 2017. AP Program Results: Class of 2017. https://reports.collegeboard.org/ap- program-results/class-2017-data.

Council on Community Pediatrics. 2016. "Poverty and Child Health in the United States."

Donner, A., Klar, N. 1996. "Statistical considerations in the design and analysis of community intervention trials." Journal of Clinical Epidemiology, 49 435-9.

Dougherty, C., Mellor, L., and Jian, S. 2006. "The Relationship between Advanced Placement and College Graduation." 2005 AP Study Series, Report 1.

Favero, Nathan. 2017. "The Geography of Teacher Shortages." US News and World Report.

Ferenstein, Gregory. 2015. "CEOs Of Silicon Valley's Top Firms Are Often Non-White, Immigrants, Or Women (In 1 Graph)." Forbes.

Fields, A., Holder, K., and Burd, CharlynnKelly Ann Holder. 2016. Life Off the Highway: A Snapshot of Rural America. US Census Bureau. https://www.census.gov/newsroom/blogs/random-samplings/2016/12/life_off_the_highway.html
.

Holztman, D. 2010. The Advanced Placement Teacher Training Incentive Program: Estimating the Impact of an Incentive and Training Program on AP Taking and Passing. American Institutes for Research.

Jackson, C.K. 2014. "Do college-preparatory programs improve long-term outcomes?" Economic Inquiry, 52(1), 72-99.

Jackson, C.K. 2010. "A stitch in time: The effects of a novel incentive-based high school." Unpublished paper submitted for publication.

Jackson, R. 2007. National Evaluation of Early Reading First. Washington, DC: Institute of Education Sciences.

Langdon, D., McKittrick, G., Beede, D., Khan, B., and Domn, M. 2012. "STEM: Good Jobs Now and for the Future." http://www.esa.doc.gov/sites/default/files/stemfinaly july14_1.pdf.

McCarthy, Niall. 2017. "The Countries With The Most STEM Graduates." February 2. https://www.forbes.com/sites/niallmccarthy/2017/02/02/the-countries-with-the-most- stem-graduates-infographic/.

National Center for Education Statistics. 2018. https://nces.ed.gov/ccd/districtsearch/.

National Center for Education Statistics. 2017. Fast Facts: Back to School Statistics.

https://nces.ed.gov/fastfacts/display.asp?id=372.

National Center for Education Statistics. 2016. Public elementary and secondary school enrollment, number of schools, and other selected characteristics, by locale: Fall 2011 through fall 2014. https://nces.ed.gov/programs/digest/d16/tables/dt16_214.40.asp.

National Center for Women and Information Technology. 2011. "Computing Education and Future Jobs: National, State & Congressional District Data."

Pew Research Center. 2014. "The Rising Cost of Not Going to College." http://www.pewsocialtrends.org/2014/02/11/the-rising-cost-of-not-going-to-college/.

Player, D. 2015. "The Supply and Demand for Rural Teachers."

Rubin, D.B. 2007. "The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials." Statistical Medicine (26), 20-36.

Sherman, D., Darwin, M., Song, M., Li, Y., Stachel, S. 2015. First-year Impacts of the National Math and Science Initiative's Advanced Placement Training and Incentive Program on Student Outcomes. Chicago, IL: American Educational Research Association.

Sherman, D., Darwin, M., Song, M., Li, Y., Statchel, S. 2015. Evaluation Findings for the College Readiness Program on Student Outcomes. Washington, DC: American Educational Research Association.

Sherman, D., Darwin, M., Stachel, S. 2014. "First-year impact of the National Math and Science Initiative's College Readiness Program." International Congress for School Effectiveness and Improvement. Cincinatti, OH.

Sherman, D., Yibing, Li., Darwin, M., Taylor, S., Song, M. 2017. Final Report on the Impacts of the National Math and Science Initiative's College Readiness Program on HIgh School Students' Outcomes. American Institutes for Research.

United States Department of Agriculture. 2017. "Rural Education at a Glance, 2017 Edition."

US Bureau of Labor Statistics. 2018. Occupational Outlook Handbook. https://www.bls.gov/ooh/computer-and-information-technology/home.htm.

US Department of Agriculture. 2017. The National Lunch Program Fact Sheet.

US Department of Agriculture. 2017. "The National School Lunch Program Fact Sheet." https://fns-prod.azureedge.net/sites/default/files/cn/NSLPFactSheet.pdf.

US Department of Education. 2017. "Status and Trends in the Education of Racial and Ethnic Groups 2017."

US News and World Report. 2016. "The U.S. News/Raytheon STEM Index Shows America Will Have to Depend on Foreign Workers to Fill STEM Jobs." https://www.usnews.com/info/blogs/press-room/articles/2016-05-17/the-us-news- raytheon-stem-index-shows-america-will-have-to-depend-on-foreign-workers-to-fill- stem-jobs.