|  |
| --- |
| Advancing STEM and English AP Program Evaluation—Year 8  AP Exam Taking and Passing Rates, Analysis Report  Final Report |
| August 2020 |

Advancing STEM and English AP Program Evaluation—Year 8

*AP Exam Taking and Passing Rates, Analysis Report*

Prepared by the UMass Donahue Institute’s  
Applied Research & Program Evaluation Group

|  |  |  |
| --- | --- | --- |
| **Project Director**  Jeremiah Johnson, Ph.D.  Senior Research Manager  **Project Manager**  Donna Spraggon, M.Sc.  Senior Research Analyst  **Primary Analyst**  Katie Ness, M.S.  Research Analyst |  | |
| The University of Massachusetts Donahue Institute is an outreach and economic development arm of the University of Massachusetts. Established in 1971, the Institute strives to connect its clients with the resources of the University, bridging theory and innovation with real world public and private sector applications. For more information: [www.donahue.umassp.edu](http://www.donahue.umassp.edu/).  The Institute’s Applied Research & Program Evaluation (ARPE) group enables education, public health, and human service organizations to make data-driven decisions to enhance program quality and capacity.  Specializing in rigorous and innovative social science research methods, ARPE works closely with federal, state, and local agencies, quasi-public agencies, and both non-profit and for-profit organizations to support programmatic and system-wide decision making. | |

Contents

[Acknowledgments 2](#_Toc49550270)

[Executive Summary 3](#_Toc49550271)

[Key Findings 4](#_Toc49550272)

[Introduction 6](#_Toc49550273)

[Evaluation Design 8](#_Toc49550274)

[Data and Data Analysis 11](#_Toc49550275)

[Methods for Quasi-Experimental Analysis 13](#_Toc49550276)

[CITS Analysis 13](#_Toc49550277)

[DID Analysis 14](#_Toc49550278)

[Propensity Score Weighting for CITS and DID Models 15](#_Toc49550279)

[Time Intervals for CITS and DID Models 17](#_Toc49550280)

[Limitations 17](#_Toc49550281)

[Findings 19](#_Toc49550282)

[Impacts on AP Exam Taking and Passing Rates 22](#_Toc49550283)

[Summary 33](#_Toc49550284)

[Appendices 35](#_Toc49550285)

[Appendix A 36](#_Toc49550286)

[Appendix B 38](#_Toc49550287)

[Appendix C 40](#_Toc49550288)

[Appendix D 49](#_Toc49550289)

[Appendix E 58](#_Toc49550290)

# Acknowledgments

The UMass Donahue Institute extends its sincere appreciation to the many people who supported and collaborated with us on this evaluation. In particular, we want to thank personnel from the Massachusetts Department of Elementary and Secondary Education, and Mass Insight Education and Research.

# Executive Summary

The Massachusetts Department of Elementary and Secondary Education (DESE) is engaged in numerous initiatives to increase the college and career readiness of students in the Commonwealth, to reduce proficiency gaps and improve academic achievement for all population groups, and to enhance the “STEM pipeline” of students who are interested in and well prepared for postsecondary education and careers in science, technology, engineering, and mathematics (including computer science). One of these initiatives is the AP STEM and English program. As specified by DESE, the goals of the program center on promoting student achievement in Advanced Placement courses and exams, especially among historically underserved populations, and school and teacher transformation related to the delivery and sustainment of AP programming.

To meet these program goals and track efforts to improve student achievement, DESE contracted with Mass Insight Education and Research (Mi) as a vendor to implement tasks and responsibilities aligned with the purposes of the program. Mi aims to implement four key tasks in partner schools that participate in the statewide program:

1. Increase participation and improve performance in AP science, mathematics, and ELA courses and on exams, with a focus on historically underserved students.
2. Increase the effectiveness of AP science, mathematics, and ELA teachers.
3. Increase the number of new and/or additional AP science, mathematics, and ELA courses offered by districts and schools in the Commonwealth.
4. Develop collaborations with other existing and/or newly established AP initiatives or organizations to build a robust and collaborative support system for historically underserved students, their parent(s)/guardian(s), and teachers.

In their work to complete these tasks, Mi is responsible for a variety of activities falling in three main tiers of assistance: 1) teacher supports, 2) student supports, and 3) school supports. Considering these three overarching types of support offered by Mi, the Year 8 evaluation study focused on measuring the outcomes that occurred as a result of teachers, students, and/or schools receiving Mi support. More specifically, this year’s evaluation focused on three program outcomes, which note that schools participating in Mi’s program should show increased AP course availability, more students taking AP exams and courses, and more students scoring a 3 or better on AP exams.

To measure these three program outcomes, the Year 8 study will include a series of reports that will provide descriptive information and measure program outcomes on AP course availability, AP course participation and passing rates, and AP exam participation and passing rates for AP STEM and English courses:

* AP Course Availability Report
* AP Course Participation and Passing Rates, Descriptive Report
* AP Course Participation and Passing Rates, Analysis Report
* AP Exam Participation and Passing Rates, Descriptive Report
* AP Exam Participation and Passing Rates, Analysis Report

DESE contracted with the University of Massachusetts Donahue Institute (UMDI) to conduct the multiyear evaluation of the Mi program. UMDI previously submitted interim and final evaluation reports for Years 1–7. This report for the Year 8 evaluation provides a summary of the quasi-experimental analyses conducted to compare AP exam participation and passing rates of students at participating schools to those of non-participating schools, and will aim to respond to the following questions:

1. Has there been a change in the percentage of students—particularly underrepresented and economically disadvantaged students—taking AP science, mathematics, and ELA exams at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students taking AP science, mathematics, and ELA exams)?
2. Has there been a change in the percentage students—particularly underrepresented and economically disadvantaged students—passing AP science, mathematics, and ELA exams at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students passing AP science, mathematics, and ELA exams)?
3. What portion of schools experience marked increases in AP exam participation during their first year of program participation?

UMDI collected several secondary data sources from DESE to address these questions.

**Key Findings**

In brief, findings in this report indicate that:

* The AP STEM and English program generally had a positive effect on student AP exam taking and passing rates during schools’ first year of participation, though results were mixed by subject and student group. For nearly all student groups and subjects, the intervention did not impact the rate of AP exam taking or passing after the first year of program participation.
* The AP STEM and English program had a positive effect on AP exam taking and passing rates for ELA, math, and any (ELA, math, or science) during the first year of the intervention.
* The percentage of students *taking* ELA, math, and any (ELA, math, or science) exams increased 5.15 percentage points, 2.17 percentage points, and 6.39 percentage points (respectively) more at participating schools than at similar non-participating schools during the first year of participation.
* The percentage of students *passing* math and science exams increased 0.72 percentage points and 0.93 percentage points (respectively) more at participating schools than at similar non-participating schools during the first year of participation.
* The effect of the AP STEM and English program on AP exam taking and passing rates for historically underserved populations was mixed.
* The program had a positive effect on AP exam taking and passing rates for ECODIS students, as well as for non-ECODIS students. The rate of *taking* any AP exam increased more for non-ECODIS students (5.43 percentage points) than it did for ECODIS students (2.81 percentage points). Similarly, the rate of *passing* any AP exam increased more for non-ECODIS students (2.34 percentage points) than it did for ECODIS students (1.74 percentage points).
  + The program had no effect on AP exam taking rates for English language learners.
  + The program had no effect on AP exam passing rates for African American / Black students, Hispanic/Latino students, Multi-Race and Non-Hispanic/Latino students, English language learners, or students with disabilities.
* Nearly all participating schools experienced an increase in the rate of AP exam participation during their first year of program participation. The proportion of schools that increased AP exam participation for specific academic discipline groupings during their first year of participation varied by cohort.
* Future analyses will need to account for the shrinking comparison group and growing dissimilarity between the participating schools and non-participating schools. Alternate—and perhaps less rigorous—weighting and modeling procedures should be considered.

# Introduction

The Massachusetts Department of Elementary and Secondary Education (DESE) is engaged in numerous initiatives to increase the college and career readiness of students in the Commonwealth, to reduce proficiency gaps and improve academic achievement for all student groups, and to enhance the “STEM pipeline” of students who are interested in and well prepared for postsecondary education and careers in science, technology, engineering, and mathematics (including computer science).

One of these initiatives is the AP STEM and English program (hereafter, “the program”). As specified by DESE, the goals[[1]](#footnote-1) of the program are to:

1. Increase AP science, mathematics, and ELA course availability, particularly at schools with limited AP science, mathematics, and ELA offerings and high percentages of historically underserved students,
2. Increase historically underserved students’ participation in AP science, mathematics, and ELA courses, such that the demographics of these courses better reflect the diversity of the student population of the school and district,
3. Increase student performance in AP science, mathematics, and ELA courses,
4. Increase the number of students taking AP exams, particularly historically underserved students,
5. Increase the number of students scoring a 3 or higher on AP exams, particularly underrepresented minority students and students who are economically disadvantaged,
6. Increase readiness for college-level study in science, mathematics, and ELA fields, and
7. Improve science, mathematics, and ELA teacher effectiveness, including content knowledge and pedagogical skills.

To meet these program goals and track efforts to improve student achievement, DESE contracted with Mass Insight Education and Research (Mi) as a vendor to implement tasks and responsibilities aligned with the purposes of the program. Mi has administered the program since the initiative’s inception in 2007, including under the program’s previous name, Mass Math + Science Initiative’s (MMSI) “Advanced Placement Training and Awards Program”.

Mi’s first cohort of the program was comprised of eight schools during the 2008–09 academic year. The program welcomed its 12th cohort of schools during the 2019–20 academic year. While Mi has continued to evolve their program over time to continue meeting the needs of participating schools, the key tasks to be implemented by the statewide program have been consistent over time. Namely, Mi aims to do the following in each participating school:

1. Increase participation and improve performance in AP science, mathematics, and ELA courses and on exams, with a focus on historically underserved students.
2. Increase the effectiveness of AP science, mathematics, and ELA teachers.
3. Increase the number of new and/or additional AP science, mathematics, and ELA courses offered by districts and schools in the Commonwealth.
4. Develop collaborations with other existing and/or newly established AP initiatives or organizations to build a robust and collaborative support system for historically underserved students, their parent(s)/guardian(s), and teachers.

In their work to complete these tasks, Mi is responsible for a variety of activities that fall into three main tiers of assistance: 1) teacher supports, 2) student supports, and 3) school supports. Specific examples of these forms of assistance include, but are not limited to, maintaining partnerships with schools with high percentages of minority and economically disadvantaged students, encouraging recruitment of minority and economically disadvantaged students into AP science and mathematics classes, providing exam fee subsidies to economically disadvantaged students, and supporting professional development for STEM AP teachers.

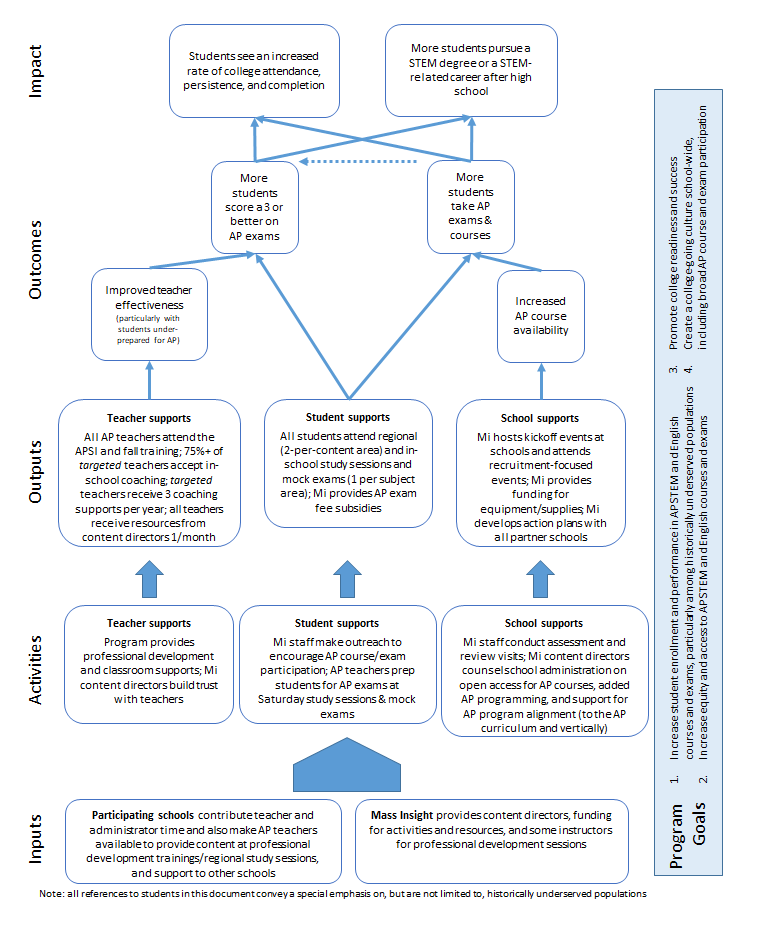
DESE contracted with the University of Massachusetts Donahue Institute (UMDI) to conduct the multiyear evaluation of the Mi program. UMDI previously submitted interim and final evaluation reports for Years 1–7. This report for the Year 8 evaluation provides a summary of the quasi-experimental analyses conducted to compare AP exam participation and passing rates of students at participating schools to those of non-participating schools.

This report is organized into five main sections. The first three sections—Evaluation Design, Data and Data Analysis, and Methods for Quasi-Experimental Analysis—provide an overarching description of the Year 8 analysis, and summarize the methods of analysis for the quasi-experimental analysis. The fourth section—Findings—presents the key results from the quasi-experimental analysis. A final section summarizes this final part of the Year 8 work.

# Evaluation Design

The Year 8 evaluation was designed to expand upon the work completed during two recent years—to be specific, Year 5 and Year 7. The Year 5 evaluation utilized a quasi-experimental design to assess outcomes for schools participating in the program. This quasi-experimental design was repeated for the Year 8 evaluation with additional years of data—although the design was modified to utilize the most robust statistical analysis possible with the available data. The outcomes measured by the quasi-experimental design were from the program’s logic model (Figure 1), which was updated during Year 7.

**Figure 1. AP STEM and English Program Logic Model**



The Year 8 evaluation study will focus on measuring three of the four outcomes included in the logic model: increased AP course availability, more students take AP exams and courses, and more students score a 3 or better on AP exams. Research questions intended to map on to these three outcomes were developed and finalized by DESE, Mi, and UMDI. These research questions fall into four categories and are as follows:

AP Course Availability

1. How many AP science, mathematics, and ELA courses have been available in participating schools over time, particularly in schools with limited offerings for economically disadvantaged and underrepresented minority students?
2. Has there been a change in the availability of AP science, mathematics, and ELA courses in participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased availability of AP science, mathematics, and ELA courses)?

AP Course Participation and Passing

1. Has there been a change in the percentage of students—particularly underrepresented and economically disadvantaged students—taking AP science, mathematics, and ELA courses at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students taking AP science, mathematics, and ELA courses)?
2. Has there been a change in the percentage of students—particularly underrepresented and economically disadvantaged students—passing AP science, mathematics, and ELA courses at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students passing AP science, mathematics, and ELA courses)?

AP Exam Participation and Passing

1. Has there been a change in the percentage of students—particularly underrepresented and economically disadvantaged students—taking AP science, mathematics, and ELA exams at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students taking AP science, mathematics, and ELA exams)?
2. Has there been a change in the percentage students—particularly underrepresented and economically disadvantaged students—passing AP science, mathematics, and ELA exams at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students passing AP science, mathematics, and ELA exams)?

Summary Questions

1. What can be learned—through descriptive analysis—about trends in AP exam taking?
2. What portion of schools experience marked increases in AP exam participation during their first year of program participation?

To answer these research questions, the Year 8 study will deliver a series of reports that will provide descriptive information and measure program outcomes on (1) AP course availability, (2) AP course participation and passing rates, and (3) AP exam participation and passing rates for AP STEM and English courses. With the exception of the report on AP Course Availability, one report with descriptive statistics and one report with the results from the quasi-experimental analysis will be written for each of these outcome areas, for a total of five reports:

* AP Course Availability Report
* AP Course Participation and Passing Rates, Descriptive Report
* AP Course Participation and Passing Rates, Analysis Report
* AP Exam Participation and Passing Rates, Descriptive Report
* AP Exam Participation and Passing Rates, Analysis Report

This fifth and final report presents results from quasi-experimental analyses completed to address research questions 4, 5, and 6b on AP exam participation and passing:

1. Has there been a change in the percentage of students—particularly underrepresented and economically disadvantaged students—taking AP science, mathematics, and ELA exams at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students taking AP science, mathematics, and ELA exams)?
2. Has there been a change in the percentage students—particularly underrepresented and economically disadvantaged students—passing AP science, mathematics, and ELA exams at participating schools compared to similar non-participating schools? If change is observed, is it positive (i.e., increased percentage of students passing AP science, mathematics, and ELA exams)?
3. What portion of schools experience marked increases in AP exam participation during their first year of program participation?

Information on the data included in these analyses, and how the data was analyzed follows in the next section.

# Data and Data Analysis

This report includes a summary of analyses comparing (1) the rates with which students at participating schools *take* AP ELA, math, and/or science exams to those rates at similar non-participating schools, and (2) the rates with which students at participating schools *pass* AP ELA, math, and/or science exams to those rates at similar non-participating schools.

These analyses are based on AP exam data provided by DESE from SY07 to SY19.[[2]](#footnote-2) Data were merged with corresponding SIMS data in order to identify key demographic information for participating students—where participating students were those in grades 9–12, and were enrolled in schools that are part of the AP STEM and English program. The newly merged data file included key information on which students were taking ELA, math, and science AP exams and which students were considered AP exam passers. In order to be considered an exam taker a student must have taken one or more AP ELA, math, or science exam. In order to be considered an exam passer a student must have scored a three or better on at least one of their AP ELA, math, or science exams.

The quasi-experimental approaches used to compare AP exam taking and passing rates at participating schools to those of similar non-participating schools were comparative interrupted time series (CITS) analyses and difference-in-difference (DID) analyses. The CITS analyses—which are described in further detail in the next section—included schools that started the intervention between SY11 and SY17 (Cohorts III–IX). Earlier and later cohorts were not included in this analysis because the years of data required to complete a CITS model were not available. In total, 86 schools were included in the treatment group (participating schools) and 222 schools were considered for inclusion in the comparison group (non-participating schools).[[3]](#footnote-3) The actual number of schools included in the models varied by subject and by outcome (taking/passing).

Additionally, DID analyses—which are described in further detail in the next section—were conducted with schools that started the intervention during SY16 and SY17 (Cohorts VIII–IX). Earlier and later cohorts for the DID analyses are not included because the years of data required to complete DID analyses were not available. The DID analyses examined the change in the percentage of students—that are economically disadvantaged—taking and passing AP ELA, math, and science exams. Additionally, DID analyses were conducted to assess the impact of the program on AP exam taking and passing rates for all students. In total, 18 schools were included in the treatment group for DID analyses, and 222 schools were considered for inclusion in the comparison group. The actual number of schools included in the models varied by subject and by outcome (taking/passing).

This report summarizes the results of quasi-experimental analyses that compared the AP exam taking and passing rates at participating AP STEM and English schools to those at non-participating schools. Quantitative results are presented by subject and student group, and examine the impact of the program on the following:

* the percentage of students taking and passing at least one AP ELA, math, or science exam the year participation began;
* the annual change in the percentage of students taking and passing at least one AP ELA, math, or science exam; as well as,
* the percentage of economically disadvantaged students taking and passing at least one AP ELA, math, or science exam two years after participation began.

A summary of results is presented in the “Findings” section. Full model results for all CITS and DID analyses are provided in Appendices C and D.

# Methods for Quasi-Experimental Analysis

The AP STEM and English program is a school-level intervention. As such, analyses to assess the program’s impact on AP exam taking and exam passing rates were conducted at the school level, comparing participating AP STEM and English program schools (treatment schools) to similar schools that did not participate in the program (comparison schools).

The impacts of the AP STEM and English program were assessed for both AP exam taking rates and for AP exam passing rates. Rates were calculated as the number of students taking/passing an AP ELA, math, or science exam divided by the total number of enrolled high school students in a school.[[4]](#footnote-4),[[5]](#footnote-5)

Program impact on AP exam taking and passing was assessed using both comparative interrupted time series (CITS) analysis and difference-in-difference (DID) analysis. The CITS analysis was the primary method of analysis conducted, whenever possible, because of its robustness. CITS analyses allowed for a comparison of treatment and control schools, and accounted for baseline trends in the outcomes of interest and other covariates of interest.

## CITS Analysis

Differences in treatment and comparison schools were assessed using a comparative interrupted time series design. In this design, AP exam taking/passing rates in treatment and comparison schools are observed across multiple school years before and after the introduction of the AP STEM and English program. The AP STEM and English program is intended to “interrupt” the level of AP exam taking/passing rates—and/or the trend (i.e., the change over time) in AP exam taking/passing rates—that would have been observed in the absence of the intervention. Using both AP STEM and English program schools and comparison schools is what makes the interrupted time series “comparative,” and this enables stronger inferences about what AP exam taking/passing rates, and trends, would have been observed in absence of the AP STEM and English program.

The CITS models assessed the effects of the AP STEM and English program on four academic groupings: (1) taking/passing rates for any AP ELA, math, or science exam, (2) taking/passing rates for AP ELA exams, (3) taking/passing rates for AP math exams, and (4) taking/passing rates for AP science exams. For each of these four academic groupings, program impact on AP exam taking/passing rates were assessed as an aggregate—across all students in AP STEM and English schools—and individually by various student groups. Finally, for each set of analyses previously mentioned, two sets of models were conducted; one to assess the impact of the program on AP exam taking rates and the other assessing program impact on AP exam passing rates. Each set of analyses included 12 models for each of the four academic groupings, yielding 48 models for each of the outcome measures; for a total of 96 CITS models.

1. CITS analyses on all students: Impacts on all students in all AP STEM and English schools. (Four academic discipline groupings and two outcomes measured yielded eight CITS models.)
2. CITS analyses on student groups: Impacts on groups of students in all AP STEM and English schools. Student groups that were assessed were female, male, Asian, African American / Black, Hispanic/Latino, Multi-Race and Non-Hispanic/Latino, White, English language learners (ELL), non-English language learners (non-ELL), students with disabilities (SWD), students without disabilities (non-SWD).[[6]](#footnote-6) (Eleven student groups, four academic discipline groupings, and two outcomes yielded 88 CITS models.)

## DID Analysis

Differences in treatment and comparison schools were also assessed using a difference-in-difference (DID) design. This model calculates the effect of a treatment on an outcome by comparing the average change over time in the outcome variable for the treatment group, compared to the average change over time for the control group. For this study, AP exam taking and passing rates were observed one school year before and two school years after the introduction of the AP STEM and English program to see if the differences in AP exam taking and passing rates, two years after participation, were significantly different from differences in the same rates at similar comparison schools. Using both AP STEM and English and comparison schools enables stronger inferences about what AP exam taking and passing rates would have been observed in the absence of the AP STEM and English program.

The DID models assessed the effects of the AP STEM and English program on four academic groupings: (1) taking/passing rates for any ELA, math, or science AP exam, (2) taking/passing rates for ELA AP exams, (3) taking/passing rates for math AP exams, and (4) taking/passing rates for science AP exams. For each of these four academic groupings, program impact on AP exam taking and passing rates were assessed for students deemed economically disadvantaged (ECODIS), as a measure of assessing program impact on students from low-income households. ECODIS was assessed as an aggregate—across all students in AP STEM and English schools—and individually by the following two student groups: economically disadvantaged and non-economically disadvantaged. Finally, for each of the analyses previously mentioned, two sets of models were conducted; one to assess the impact of the program on AP exam taking rates and the other assessing program impact on AP exam passing rates. Each set of analyses included three models for each of the four academic groupings, yielding 12 models for each of the outcome measures; for a total of 24 DID models.

1. DID analyses on all students: Impacts on all students in all AP STEM and English schools. (Four academic discipline groupings on two outcomes measured yielded eight DID models.)
2. DID analyses on student groups: Impacts on groups of students in all AP STEM and English schools. The student groups assessed were economically disadvantaged and non-economically disadvantaged. (Two student groups, four academic discipline groupings, and two outcomes yielded 16 DID models.)

Program impact on AP exam taking/passing rates was determined through a procedure called propensity score weighting.

## Propensity Score Weighting for CITS and DID Models

The AP STEM and English program did not utilize random assignment because each school was selected by Mi to participate based on school characteristics. Therefore, it is likely that there were differences between treatment and comparison schools prior to intervention. These differences could have represented a significant threat to the validity of the study’s findings. To reduce these differences substantially, propensity score weighting procedures were used, thereby improving the validity of the estimates of program impacts.[[7]](#footnote-7)

In essence, propensity score weighting is used to approximate the results of random assignment by reducing multiple covariates (e.g., race/ethnicity, gender, AP exam taking/passing rates prior to the AP STEM and English program) to a single score called a propensity score. A propensity score was calculated for each treatment and comparison school that described the likelihood of that school participating in the AP STEM and English program. Weighting procedures were then applied to balance propensity scores for treatment and comparison schools. Propensity scores allowed the research team to generate estimates of the average treatment effect for the treated (ATT) population. This approach is typical for quasi-experimental studies that try to assess the impact of a particular program, such as the AP STEM and English program.

Covariates used in the propensity score weighting procedure included pre-intervention rates of AP exam taking/passing (from the four school years prior to intervention for CITS analyses and one year prior to intervention for DID analyses), gender, race/ethnicity, ELL status, special education status, and average school MCAS CPI (by subject, as appropriate for each analysis). In addition, for DID analyses, the propensity score weighting procedure also included the percentage of ECODIS students one year prior to intervention, as applicable. Once weights were assigned, the balance of the covariate distributions between treatment and comparison schools was assessed in terms of standardized bias. For this study, we considered a covariate to be balanced if the standardized bias was less than 0.25. Although there is no universal criterion for assessing precisely when balance has been achieved, 0.25 is commonly used.[[8]](#footnote-8)

To assess each CITS and DID model’s potential for producing findings with a high likelihood of validity, the balance of covariates after weighting was considered. When propensity score weighting was completely successful, it yielded a comparison group that met the balance criteria (i.e., standardized bias less than 0.25) for all covariates. Models in which all covariates met this criterion after weighting were designated as “fully balanced.” Models that were not fully balanced were assessed to see if more than half of the covariates used in the weighting procedure achieved a standardized bias of less than 0.25 after weighting. Models that achieved this criterion were designated as “partially balanced.” Results are not reported for models that did not achieve full or partial balance because no adequately matched comparison group could be identified. Of the 96 CITS models assessed, 25 were fully balanced after weighting, 42 were partially balanced, and 28 were not balanced. Of the 24 DID models assessed, 22 were fully balanced after weighting, and two were partially balanced.

For the 28 models that were not balanced, the data consistently showed that differences in the proportion of students in the treatment and comparison schools taking and passing AP exams existed prior to the intervention. Further, the proportion of students taking and passing AP exams was lower in treatment schools than in comparison schools. Over time, schools selected for intervention appear to be those with lower pre-intervention exam taking and passing rates.

In models where full balance or partial balance was not achieved, an alternate weighting model—including only pre-intervention rates of AP exam taking/passing from the four school years prior to intervention as covariates—was applied. This alternate weighting procedure was used because it allowed us to weight the control sample so that the trend in AP exam taking and passing rates prior to intervention was similar to that of the treatment group. Of the 28 CITS models that were not balanced after the original weighting procedure, 24 were fully balanced, two were partially balanced, and two remained not balanced after the alternate weighting procedure was competed. CITS analyses that relied on samples constructed using this alternate weighting procedure are presented with a footnote. These results should be interpreted with caution, as the surety of the comparability between the treatment and comparison groups is not guaranteed.

Over time—as additional schools join the AP STEM and English program—fewer schools will be eligible for inclusion in the comparison group. Because schools are carefully selected for participation, and because propensity score matching and weighting procedures are dependent on the comparability of treatment and comparison samples, it is likely that it will become increasingly difficult to identify comparison schools that are similar to those participating in the intervention. Alternative—and likely less rigorous—weighting and/or modeling procedures will need to be considered for future analysis.

Even if samples met the criteria described above for full balance or partial balance, it may be determined from CITS or DID analysis that baseline pre-intervention exam taking/passing rates—or trends in rates—differed for AP ELA, math, and science exams.[[9]](#footnote-9) While such differences raise some concerns about the ability to draw causal inferences about the relationship between the AP STEM and English program and AP exam taking/passing rates, the full or partial balance achieved via the propensity score weighting (either full weighting or alternate weighting) provides evidence of substantial similarity between AP STEM and English participant schools and comparison schools.

## Time Intervals for CITS and DID Models

It is important to note that both CITS and DID analyses were dependent on the number of years for which data on AP exam taking and passing rates were available. The time intervals for assessing impacts were based on the number of years between a given AP exam taking rate or AP exam passing rate being available and when a school joined the AP STEM and English program. For CITS analyses, only cohorts for which four years of pre-intervention data and two years of post-intervention data (seven total years of data) were available were eligible for inclusion in this analysis. Data from SY07 to SY19 were available for CITS analyses. As such, the cohorts for which the necessary data was available to conduct CITS analyses included Cohort III through Cohort IX. Cohort III joined the AP STEM and English program in SY11, while Cohort IX joined in SY17. For DID analyses, only cohorts for which one year of pre-intervention data and two years post-intervention data (two total years of data) were available were eligible for inclusion in this analysis. Data from SY15 to SY19 was available for DID analyses, as DESE began using ECODIS as a proxy for low-income in SY15. As such, the cohorts for which the necessary data were available to conduct DID analyses for ECODIS included Cohorts VIII and IX. Cohort VIII joined the AP STEM and English program in SY16 and, again, Cohort IX joined in SY17.

## Limitations

In order for a school to be included in CITS analyses, we required that the school have seven years of relevant data; four years of baseline data, and three years of data after program implementation. Because of the CITS data requirements, and because the measure used to identify students from low-income families changed during the study period,[[10]](#footnote-10) the research team could not include a variable that accounted for differences in schools based on the proportion of students who were from low-income families in CITS analyses. The research team understands that level of income is an important predictor of student success, and that students from low-income families are a student group of interest for this intervention. In order to assess the impact of the intervention on students from low-income families—and to account for income status in the overall assessment of program impact—the research team completed a series of DID analyses. While DID analyses do not account for baseline trends in outcomes of interest, the data requirements allowed for the inclusion of ECODIS as a covariate, identifying the proportion of schools’ students that were from low-income families―after the start of SY15.

With the available data for ECODIS analyses being limited to SY15–SY19, combined with the data requirements for DID analysis, the number of eligible cohorts and hence sample size of schools included in the DID analyses was relatively small. As such, the results of all DID analyses should be interpreted with caution.

As noted before, in order to meet balancing criteria an alternate, less comprehensive, weighting procedure was applied to samples in which the original weighting procedure did not result in full or partial balance. These results should be interpreted with caution.

The “Findings” section below summarizes program impact analysis results, indicating which models were fully or partially balanced, and which models showed a significant difference in the rates for AP exam taking and AP exam passing. Additionally, the CITS analyses also indicate which models showed a significant change in the trend (i.e., change over time) of AP exam taking and passing rates after joining the program.

# Findings

This section presents (1) the findings related to the percentage of students—particularly those typically underrepresented in advanced placement courses—taking AP exams and passing AP exams (research questions 4 and 5); and (2) the findings related to the portion of schools experiencing marked increases in AP exam participation during their first year of program participation (research question 6b). Findings are organized in three subsections. The first subsection summarizes the results of the quasi-experimental analysis comparing AP exam taking and passing rates of all students at schools participating in the AP STEM and English program to similar non-participating schools. The second subsection summarizes the results of the quasi-experimental analysis comparing AP exam taking and passing rates of student groups of interest at schools participating in the AP STEM and English program to similar non-participating schools. The third subsection summarizes the results from the descriptive analysis, which examined the portion of participating schools experiencing increases in AP exam taking rates during the first year of intervention―first by academic discipline groupings (ELA, math, or science; ELA; math; and science) and then by number of AP exams.

When interpreting the findings presented within this section, it is important to consider several notes about AP exam taking and passing rates, including the following:

1. Students at participating schools were considered an AP exam taker if they took one or more AP ELA, math, or science exam. As the AP STEM and English program aims to increase the proportion of students enrolling in AP exams, the percentage of AP STEM and English exam takers was calculated using the total number of students enrolled in 9th–12th grade at the school as the denominator—as opposed to using the total number of students taking AP exams for any subject.
2. Students at participating schools were considered an AP exam passer if they received a score 3 or better for at least one of the AP STEM and English exams they took. The percentage of AP STEM and English exam passers is defined as the number of students passing one or more AP exam divided by the number of students taking one or more AP exam.

Full results from all analyses is located in the appendices as follows: exam taking models (Appendix C), exam passing models (Appendix D), and portion of schools experiencing increases in AP exam participation (Appendix E).

|  |
| --- |
| **Summary of Key Findings** |
| * The AP STEM and English program generally had a positive effect on student AP exam taking and passing rates during schools’ first year of participation, though results were mixed by subject and student group. For nearly all student groups and subjects, the intervention did not impact the rate of AP exam taking or passing after the first year of program participation. * The AP STEM and English program had a positive effect on AP exam taking and passing rates for ELA, math, and any (ELA, math, or science) during the first year of the intervention. * The percentage of students *taking* ELA, math, and any (ELA, math, or science) exams increased 5.15 percentage points, 2.17 percentage points, and 6.39 percentage points (respectively) more at participating schools than at similar non-participating schools during the first year of participation. * The percentage of students *passing* math and science exams increased 0.72 percentage points and 0.93 percentage points (respectively) more at participating schools than at similar non-participating schools during the first year of participation. * The effect of the AP STEM and English program on AP exam taking and passing rates for historically underserved populations was mixed. * The program had a positive effect on AP exam taking and passing rates for ECODIS students, as well as for non-ECODIS students. The rate of *taking* any AP exam increased more for non-ECODIS students (5.43 percentage points) than it did for ECODIS students (2.81 percentage points). Similarly, the rate of *passing* any AP exam increased more for non-ECODIS students (2.34 percentage points) than it did for ECODIS students (1.74 percentage points).   + The program had no effect on AP exam taking rates for English language learners.   + The program had no effect on AP exam passing rates for African American / Black students, Hispanic/Latino students, Multi-Race and Non-Hispanic/Latino students, English language learners, or students with disabilities. * Nearly all participating schools experienced an increase in the rate of AP exam participation during their first year of program participation. The proportion of schools that increased AP exam participation for specific academic discipline groupings during their first year of participation varied by cohort. * Future analyses will need to account for the shrinking comparison group and growing dissimilarity between the participating schools and non-participating schools. Alternate—and perhaps less rigorous—weighting and modeling procedures should be considered. |

## Impacts on AP Exam Taking and Passing Rates

Impacts on AP exam taking and passing rates were assessed in relation to all students and student groups of interest for each of the four AP STEM and English program academic discipline groupings: ELA, math, or science; ELA; math; and science. The rate with which students at participating schools took and passed AP exams was compared to the rate with which students at similar schools that did not participate in the program (comparison schools) took and passed AP exams. Forty-eight CITS models were conducted for both AP exam taking and AP exam passing—12 models for each of the four academic discipline groupings: ELA, math, science, and ELA, math, or science—for a total of 96 models. Additionally, impacts on AP exam taking and passing rates were assessed in relation to all students and student groups of interest, related to low-income status using ECODIS, for each of the four AP STEM and English academic discipline groupings. A total of 24 DID models were run; three models for each of the four academic groupings for both outcomes (AP exam taking and AP exam passing).

Results showing the significance of the impact, if any, are presented in the body of the report. Results showing the complete models are presented in the appendices.

Statistically significant and positive program impacts were identified for 73 of the 96 CITS models (see Tables 1a and 2a). This includes 46 models that fully or partially balanced using the original weighting procedure, and 27 that balanced using the alternate weighting procedure.[[11]](#footnote-11) The tables indicate significance in relation to two aspects of AP exam taking and passing rates. The first notes the significance of any differences—between AP STEM and English participant schools and comparison schools—in the percentage of students taking/passing an AP ELA, math, and/or science exam one year after schools began participating in the AP STEM and English program. These differences are presented in the “AP Exam Taking/Passing Rate Change after One Year” column. The second aspect of change in the tables is the “Annual Change in Percent Taking/Passing Rate” column. This column indicates the significance of any differences—between AP STEM and English participant schools and comparison schools—in the annual rate of change in students’ AP exam taking/passing rate during the three years after schools began participating in the program. These changes are presented as the percentage of students taking/passing an AP exam, with a positive number indicating that participating schools had an increase in taking/passing rates.

Statistically significant and positive program impacts were identified for 19 of the 24 DID models overall, as summarized in Tables 1b and 2b. The tables indicate the level of significance in relation to the AP exam taking/passing rate change two years after participation for AP STEM and English schools when compared to similar non-participating schools. These changes are presented as the percentage of students taking/passing an AP exam, with a positive number indicating an increase in taking/passing rates. **Table 1a: Summary of AP STEM and English Impact on Exam Taking and Passing by Subject Percentage Point Change—CITS for All Students**

| **Student Group** | **Subject2** | **Exam Taking1** | | **Exam Passing1** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **AP Exam Taking Rate Change after One Year** | **Annual Change in Percent Taking Rate** | **AP Exam Passing Rate Change after One Year** | **Annual Change in Percent Passing Rate** | |
| **All students** | Any | 6.39\*\*\* ∇ | n.s. ∇ | n.r. ∇ | n.r. ∇ | |
| ELA | 5.15\*\*\* ∇ | n.s. ∇ | n.r. ∇ | n.r. ∇ | |
| Math | 2.17\*\*\* ∇ | 0.33+ ∇ | 0.72\*\*\* | n.s. | |
| Science | n.r. ∇ | n.r. ∇ | 0.93\*\*\* | 0.17+ | |
| *Note*: + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  ∇Alternate weighting was used. See methods for description.  “n.s.” means “no significant findings.” Only statistically significant results are presented.  “n.r.” means “not reported.” Findings are not reported because either the models did not converge and therefore produced no results or the samples of participating and non-participating schools were not balanced or partially balanced.  1 Change in percentage points of students taking/passing an AP exam at a school. A positive number indicates an increase in the percentage taking/passing.  2 “Any” refers to any AP ELA, math, or science exam. | | | | | |

**Table 1b: Summary of AP STEM and English Impact on Exam Taking and Passing by Subject Percentage Point Change—DID for All Students (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Student group** | **Subject2** | **Difference in Percentage Point Change in Taking Between Treatment and Control Groups Two Years After Participation1** | **Difference in Percentage Point Change in Passing Between Treatment and Control Groups Two Years After Participation** |
| **All students** | Any | 4.35\*\*\* | 2.11\*\*\* |
| ELA | 3.04\*\*\* | 1.24\*\* |
| Math | 2.67\*\*\* | 1.21\*\*\* |
| Science | n.s. | n.s. |
| *Note*: “n.s.” means “no significant findings.” Only statistically significant results are presented.  + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students taking/passing an AP exam at a school compared to change for control group. A positive number indicates a greater increase in the percentage taking/passing for participating schools.  2 “Any" refers to any AP ELA, math, or science exam. | | | |

**Impacts on AP exam taking and passing rates for all students**

CITS and DID analyses were conducted to assess the impact of the intervention on AP exam taking rates for the four academic groupings—any, ELA, math, or science—at participating schools. Results were similar regardless of method. Results of both CITS and DID analyses indicate that the program had a positive effect on the percentage of students taking any, ELA, or math AP exam (see Tables 1a and 1b).

Results from CITS analyses indicate that—during the first year of the intervention—the percentage of students taking any AP (ELA, math, or science) exam increased an additional 6.39 percentage points at participating schools over those in comparison schools. Similarly, DID analyses indicate that any AP (ELA, math, or science) exam taking rates at participating schools increased an additional 4.35 percentage points in over those in comparison schools.

CITS and DID results suggest that the increase observed in AP ELA exam taking rates was higher than that observed for AP math exam taking rates. Further, CITS analyses indicate that AP ELA exam taking rates among participating schools increased an additional 5.15 percentage points during the first year of the intervention, when compared to comparison schools. CITS analyses for AP math exams indicate that AP exam taking rates among participating schools increased 2.17 percentage points more at participating schools during the first year of the intervention than they did at comparison schools. Math exam taking rates may have continued to increase over time for participating schools, as post intervention trends showed marginally significant (*p* < 0.1) differences between AP STEM and English participant schools and comparison schools.

Results of DID analyses were similar to the results of CITS analyses, and provide additional evidence that the intervention had a positive effect on AP ELA, AP math, and any AP exam taking. Two years after schools began participating in the program, taking rates for AP STEM and English schools were an additional 3.04 percentage points higher for ELA and 2.67 percentage points higher for math.

Both CITS and DID results suggest that the AP STEM and English program did not have a statistically significant effect on the percentage of students taking an AP science exam.

Results for AP exam passing were dissimilar across methods. CITS analyses indicate significant, additional increases in the percentage of students passing an AP math or science exam at participating schools—in comparison to the percentage of students passing AP math or science exams at comparison schools. AP math exam passing rates increased an additional 0.72 percentage points for participating schools during the first year of the intervention, while AP science exam passing rates increased 0.93 percentage points. Results suggest that increases in AP science exam passing may be sustained over time, as the result was marginally significant (*p* < 0.1). DID analyses indicate that the percentage of students passing AP exams increased more at participating schools than at comparison schools for ELA, math, and any (ELA, math, or science), but not for science. These increases were approximately one to two percentage points.

**Table 2a: Summary of AP STEM and English Impact on Exam Taking and Passing by Student Group Percentage Point Change—CITS for Student Groups**

| **Student Group** | **Subject2** | **Exam Taking1** | | **Exam Passing1** | |
| --- | --- | --- | --- | --- | --- |
| **AP Exam Taking Rate Change after One Year** | **Annual Change in Percent Taking Rate** | **AP Exam Passing Rate Change after One Year** | **Annual Change in Percent Passing Rate** |
| **Female students** | Any | 7.63\*\*\* ∇ | n.s. ∇ | 2.73\*\*\* ∇ | n.s. ∇ |
| ELA | 6.49\*\*\* ∇ | n.s. ∇ | 2.23\*\*\* ∇ | n.s. ∇ |
| Math | 2.43\*\*\* ∇ | n.s. ∇ | 0.89\*\*\* ∇ | n.s. ∇ |
| Science | 2.19\*\* | 0.85+ | 0.93\*\*\* | 0.25+ |
| **Male students** | Any | 4.82\*\*\* ∇ | n.s. ∇ | 1.92\*\*\* ∇ | n.s. ∇ |
| ELA | 3.35\*\*\* ∇ | n.s. ∇ | 1.60\*\*\* ∇ | n.s. ∇ |
| Math | 1.85\*\*\* ∇ | n.s. ∇ | 0.56\*\* | 0.25\* |
| Science | 2.17\*\*\* ∇ | n.s. ∇ | 0.90\*\*\* | n.s. |
| **African American / Black students** | Any | 4.08\*\* | n.s. | n.s. | 0.58+ |
| ELA | 2.43\* | n.s. | n.s. | n.s. |
| Math | n.r. | n.r. | n.s. | n.s. |
| Science | 1.96\* | n.s. | n.s. | 0.38+ |
| **Asian students** | Any | 10.97\*\*\* | n.s. | 3.43\* | n.s. |
| ELA | 8.87\*\*\* | n.s. | 4.42\*\* | n.s. |
| Math | 5.59\*\* | n.s. | n.s. | n.s. |
| Science | 6.12\*\* | n.s. | n.s. | n.s. |
| **Hispanic/Latino students** | Any | 4.41\*\*\* | n.s. | n.s. | n.s. |
| ELA | 3.39\*\*\* | n.s. | n.s. | n.s. |
| Math | 1.25\*\* | n.s. | n.s. | n.s. |
| Science | 1.72\* | n.s. | n.s. | n.s. |
| **Multi-Race and Non-Hispanic/Latino students** | Any | n.s. | n.s. | n.s. | n.s. |
| ELA | 2.92+ | 2.36\* | n.s. | n.s. |
| Math | n.s. | n.s. | n.s. | n.s. |
| Science | n.s. | n.s. | n.s. | n.s. |
| **White students** | Any | 6.13\*\*\* ∇ | 0.71+ ∇ | 2.94\*\*\* ∇ | n.s. ∇ |
| ELA | 4.43\*\*\* ∇ | n.s. ∇ | 2.06\*\*\* ∇ | n.s. ∇ |
| Math | 2.10\*\*\* | n.s. | 0.90\*\*\* | n.s. |
| Science | 2.65\*\*\* ∇ | 0.58\* ∇ | 1.12\*\*\* | 0.27+ |
| **English Language Learner students** | Any | n.s. | n.s. | n.s. | n.s. |
| ELA | n.s. | n.s. | n.r. | n.r. |
| Math | n.s. | 0.37+ | n.s. | n.s. |
| Science | n.s. | n.s. | n.s. | n.s. |
| **Non-English Language Learner students** | Any | 6.66\*\*\* | n.s. | 2.37\*\*\* | n.s. |
| ELA | 5.15\*\*\* | n.s. | 1.91\*\*\* | n.s. |
| Math | 2.28\*\*\* | n.s. | 0.73\*\*\* | n.s. |
| Science | n.s. | n.s. | 0.94\*\*\* | n.s. |
| **Students with Disabilities** | Any | 0.46\* | n.s. | n.s. | n.s. |
| ELA | 0.31+ | n.s. | n.s. | n.s. |
| Math | n.s. | n.s. | n.s. | n.s. |
| Science | 0.15+ | n.s. | n.s. | n.s. |
| **Students without Disabilities** | Any | 7.52\*\*\* ∇ | n.s. ∇ | 2.84\*\*\* ∇ | n.s. ∇ |
| ELA | 6.63\*\*\* ∇ | n.s. ∇ | 2.21\*\*\* ∇ | n.s. ∇ |
| Math | 2.48\*\*\* ∇ | n.s. ∇ | 0.83\*\*\* | n.s. |
| Science | 2.84\*\*\* | n.s. | 1.06\*\*\* | n.s. |
| *Note*: + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  ∇Alternate weighting was used. See methods for description.  “n.s.” means “no significant findings.” Only statistically significant results are presented.  “n.r.” means “not reported.” Findings are not reported because either the models did not converge and therefore produced no results or the samples of participating and non-participating schools were not balanced or partially balanced.  1 Change in percentage points of students taking/passing an AP exam at a school. A positive number indicates an increase in the percentage taking/passing.  2 “Any” refers to any AP ELA, math, or science exam. | | | | | |

**Table 2b: Summary of AP STEM and English Impact on Exam Taking and Passing by Student Group Percentage Point Change—DID for Student Groups (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Student group** | **Subject2** | **Difference in Percentage Point Change in Taking Between Treatment and Control Groups Two Years After Participation1** | **Difference in Percentage Point Change in Passing Between Treatment and Control Groups Two Years After Participation** |
| **Economically Disadvantaged** | Any | 2.81\*\* | 1.74\*\* |
| ELA | 2.46\*\* | 1.39\*\*\* |
| Math | 1.85\*\*\* | 0.75\*\* |
| Science | n.s. | n.s. |
| **Non-Economically Disadvantaged** | Any | 5.43\*\*\* | 2.34\*\*\* |
| ELA | 3.62\*\*\* | 1.19\*\* |
| Math | 2.96\*\*\* | 1.38\*\*\* |
| Science | 1.27+ | n.s. |
| *Note*: “n.s.” means “no significant findings.” Only statistically significant results are presented.  + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1 Change in percentage points of students taking/passing an AP exam at a school compared to change for control group. A positive number indicates a greater increase in the percentage taking/passing for participating schools.  2 “Any" refers to any AP ELA, math, or science exam. | | | |

**Impacts on AP exam taking and passing rates for student groups of interest.** One goal of the AP STEM and English program is to increase participation in and performance on AP ELA, math, and science exams, particularly for those students from traditionally underserved groups. To assess progress towards this goal, 11 student groups were assessed using CITS analyses—female, male, African American / Black, Asian, Hispanic/Latino, Multi-Race and Non-Hispanic/Latino, White, English language learners (ELL), non-English language learners (non-ELL), students with disabilities (SWD), and students without disabilities (non-SWD). Additionally, two student groups were assessed using DID analyses—economically disadvantaged (ECODIS) and non- economically disadvantaged (non-ECODIS). It is important to note that (1) the intervention either had a statistically significant and positive impact, or no impact, on all student groups for AP exam taking and passing rates during the first year of intervention; (2) the impact of the intervention on AP exam taking and passing rates varied by AP exam and student group; and (3) any statistically significant changes in exam taking or passing rates observed were larger for AP ELA and any AP (ELA, math, or science) than for those observed in AP math and AP science for every student group.

When significant differences between participating and comparison schools were found for AP exam *taking* rates for student groups, they ranged in magnitude from small (0.15 percentage points for AP science for students with disabilities) to large (10.97 percentage points for Asian students taking any AP—ELA, math or science—course). Overall, female, male, Asian, Hispanic/Latino, White, non-SWD, and non-ECODIS students at participating schools reported additional gains in AP exam taking rates after the first year of participation for AP ELA, AP math, AP science, and any AP (ELA, math, or science) exams, in comparison to comparison schools (see Tables 2a and 2b). With the exception of Asian, Hispanic/Latino, and non-ECODIS students, these were also the same student groups that reported significant gains in AP exam *passing* rates after year one for AP ELA, AP math, AP science, and any AP (ELA, math, or science) exams. Moreover, additional gains in AP exam passing rates after year one for AP ELA, AP math, AP science, and any AP (ELA, math, or science) exams were also found for non-ELL.

CITS analyses indicate that the intervention had little or no impact on the rate of AP exam taking or passing after the initial year of program participation. In the few cases where possible trends were detected, the results were only marginally significant (*p* < .10) or significant at the *p* < .05 level. Further results of the analysis, broken down by student groups of interest, is as follows.

**Race/Ethnicity**

Results indicate that the intervention had a positive impact on the percentage of students taking at least one AP ELA, math, science and any AP (ELA, math, or science) exam for Asian, Hispanic/Latino, and White students. Additionally, results indicate that the intervention had a statistically significant and positive impact on the percentage of African American / Black students taking at least one AP ELA, AP science, and any AP (ELA, math, or science) exam; no significant impact was detected for AP math exam for African American / Black students. Lastly, the intervention had a marginally significant (*p* < .10) and positive impact on the percentage of Multi-Race and Non-Hispanic/Latino students taking at least one AP ELA exam.

Results indicate that the intervention had a statistically significant and positive impact on the percentage of White students passing at least one AP ELA, math, science, or any AP (ELA, math, or science) exam. Results also indicate that positive impacts were found for the percentage of Asian students passing at least one AP ELA or any AP (ELA, math, or science) exam. Notably, results of CITS analyses indicate that the intervention did not have a statistically significant effect on the percentage of Hispanic/Latino and Multi-Race and Non-Hispanic/Latino students passing AP ELA, AP math, AP science, or any AP exam.

Across racial/ethnic student groups, there is evidence that the intervention had a positive impact on rates of AP exam taking one year after the intervention started, and that rates of exam taking, in general, did not continue to increase more at participating schools than at non-participating schools. Generally, gains in AP exam passing are not seen over time, with the exception of Asian students passing at least one AP ELA or any AP (ELA, math, or science) exam, and White students passing at least one AP exam (ELA, math, science, or any). After the initial year of participation, few additional gains were observed for both rates of AP exam taking and passing.

**Gender**

Results indicate that the intervention had a positive impact on the percentage of students taking and passing at least one AP ELA, math, science, or any AP (ELA, math, or science) exam for both female and male students.

Overall, any changes observed in AP exam taking and passing rates for female and male students were statistically significant and educationally meaningful. Further, gains in all subjects for both taking and passing rates were larger for female students than for male students.

**Student Groups of Interest**

Results indicate that the intervention had a positive impact on the percentage of students taking at least one AP ELA, math, science, or any AP (ELA, math, or science) exam for non-SWD, and non-ECODIS students. Similarly, results indicate that the intervention had a positive impact on the percentage of students taking at least one AP ELA, AP science, or any AP (ELA, math, or science) exam for SWD. No significant impact was detected on the percentage of SWD taking an AP math exam. The intervention also had a positive impact on the percentage of non-ELL and ECODIS students taking at least one AP ELA, math, or any AP (ELA, math, or science) exam; no significant impact was detected for the percentage of students taking an AP science exam in either of these student groups. Notably, results of CITS analyses indicate that the intervention did not affect the percentage of ELL students taking or passing AP ELA, math, science, or any AP course.

Results indicate that the intervention had a positive impact on the percentage of non-ELL and non-SWD students passing at least one AP ELA, math, science, or any AP (ELA, math, or science) exam. The intervention also had a positive impact on the percentage of ECODIS and non-ECODIS students passing at least one AP ELA, math, or any AP (ELA, math, or science) exam. Notably, results of CITS analyses indicate that the intervention did not have a statistically significant effect on the percentage of ELL and SWD passing AP ELA, math, science, or any AP exam.

The intervention did not have a statistically significant impact on the percentage of ELL students taking or passing an AP exam. Small gains were detected for SWD in AP exam taking, but no significant gains were detected for exam passing, after the first year of intervention. There were significant gains in the AP exam taking and passing rates of non-ELL and non-SWD students; however, there was no indication that AP exam taking or passing rates continued to increase in subsequent years.

Significant gains in AP exam taking and passing were observed for ECODIS students in all subjects with the exception of AP science. Significant gains in AP exam taking and passing were also observed for non-ECODIS students; however, non-ECODIS students had significant gains in exam taking rates in AP science.

**Schools experiencing marked increases in AP exam participation during their first year of program participation.** As the above CITS analyses have shown, the AP STEM and English program generally had a positive effect on AP ELA, math, and any AP (ELA, math, or science) exam taking and passing rates during the first year of the intervention. Using descriptive analysis, Figure 2 shows that nearly all participating schools experienced an increase in the rate of AP exam participation rates during their first year of participation from SY09 to SY19. The majority of participating schools experienced an increase in AP exam participation rates in all three subject areas (ELA, math, and science) from SY09 to SY16. After SY16, between SY17 and SY19, the proportion of participating schools that experienced an increased rate of AP exam participation did so in only one or two subject areas.

**Figure 2: Portion of Participating Schools Experiencing Increased AP Exam Participation During First Year of Intervention, by Number of Subject Areas**

Again, using descriptive analysis, Figure 3 shows that the rate of AP exam participation for all academic discipline groupings (ELA, math, or science; ELA; math; and science) increased for most participating schools during their first year of program participation from SY09 to SY16, and that the proportion of schools experiencing increased rates of AP exam participation by subject was generally lower and more variable from SY17 to SY19.

**Figure 3: Portion of Participating Schools Experiencing Increased AP Exam Participation During First Year of Intervention, by Academic Discipline Grouping**

# Summary

A primary goal of the AP STEM and English program is to increase student participation in and performance on AP STEM and English exams, particularly among historically underserved populations.

Results for all groups indicate that the intervention either had a positive impact, or no impact on AP exam taking and passing rates during the first year of intervention. When differences were detected, they ranged in magnitude from small (0.15 percentage points for AP science for students with disabilities) to large (10.97 percentage points for Asian students taking any AP—ELA, math or science—exam).

Nearly all participating schools experienced an increase in the rate of AP exam participation rates during their first year of participation from SY09 to SY19. Similarly, a majority of participating schools experienced an increase in AP exam participation rates in all three subject areas (ELA, math, and science) from SY09 to SY16.

Results of CITS and DID analyses varied by method. Results of CITS analyses indicate that the intervention did not have a statistically significant effect on the percentage of students passing AP ELA or any AP exam. However, DID results indicate that the program did have a statistically significant effect on the percentage of students passing AP ELA or any AP exam. Specifically, DID analyses showed that the rate of AP ELA exam passing increased 1.24 percentage points more at participating schools than at comparison schools. Similarly, DID analyses showed that the rate of passing any AP exam increased 2.11 percentage points more at participating schools than at comparison schools. Additionally, DID results indicate that the program did not have a statistically significant effect on the percentage of students taking an AP science exam.

Results also varied by student group. For example, results indicate that the intervention had a positive impact on the percentage of students taking AP ELA, math, science, and any AP (ELA, math, or science) exams for female, male, Asian, Hispanic/Latino, White, non-SWD, and non-ECODIS students. However, results also indicate that the intervention did not have a statistically significant effect on the percentage ELL students taking AP ELA, math, science, or any AP exam. Further, results indicate that the intervention had a positive impact on the percentage of students passing AP ELA, math, science, and any AP (ELA, math, or science) exams for female, male, White, non-ELL, and non-SWD students. However, results also indicate that the intervention did not have a statistically significant effect on the percentage of Hispanic/Latino, Multi-Race and Non-Hispanic/Latino students, ELL, or SWD passing AP ELA, math, science, or any AP exam.

As the AP STEM and English program has progressed, the pool of potential comparison schools has shrunk. At the same time, the dissimilarity of the schools participating in the intervention and the comparison schools grow has grown. We anticipate that these challenges will limit future analysis efforts. Future analyses should consider alternate weighting and/or modeling procedures.

Overall, findings indicate that the AP STEM and English program had a positive impact on AP exam taking and passing rates in the first year of intervention for AP ELA, math, and any AP exams. Schools participating in the AP STEM and English program experienced larger increases in the taking and passing of AP ELA exams than for AP math or science exams. Similarly, some evidence of progress being made toward the goal of increasing AP exam taking and passing rates among historically underserved populations was detected. Across the Commonwealth, more support may be needed to improve AP exam taking and passing rates for AP ELA, math, and—most notably—science exams for all student groups, and particularly among those that are underserved.

# Appendices

# Appendix A

**Modeling Procedures for Comparative Interrupted Time Series (CITS) Analyses**

For each academic discipline (i.e., ELA, math, and science), a CITS model was developed to assess the impact of the AP STEM and English program on (1) a school’s AP exam taking/passing rate one year after the program began, and (2) the trend (i.e., the slope) of the AP exam taking/passing rate during the three year period after the program began.

This procedure was used for all 96 CITS models. The following equation represents the CITS model:

*Yit = β0 + β1Timet, + β2Interventiont + β3TimetInterventiont* ***+*** *β4Participanti + β5ParticipantiTimet + β6ParticipantiInterventiont + β7ParticipantiTimetInterventiont + ui +eit*

In this model, Yit is the outcome measure for a school *i* at time *t*. *Timet* is the time in years since the start of the study. *Interventiont* is an indicator of whether or not a school was participating in the intervention at time *t*. *TimetInterventiont* is an interaction between *Timet* and *Interventiont*. *Participanti* is an indicator for a school *i* that participated in the AP STEM and English program (by academic discipline). *ParticipantiTimet* , *ParticipantiInterventiont* , and *ParticipantiTimetInterventiont* are interaction terms used in comparisons of multiple groups. Random effects were included to account for school and individual observation effects by adding a random error term for each school (ui), and individual observations (eit).

The *β0* to *β3* coefficients represent the control group. The *β4* to *β7* coefficients represent differences between the treatment and control groups. *β1* represents the slope, or trajectory of the outcome variable until the introduction of the intervention. *β2*represents the change in the level of the outcome variable that occurs in the period immediately following the introduction of the intervention. *β3*represents the difference between pre- and post-intervention slopes of the outcome. *β4* represents the difference in the level (intercept) between treatment and control prior to intervention. *β5* represents the difference in the slope between treatment and control prior to intervention. *β6* represents the impact of the AP STEM and English program on AP exam taking/passing rates at each school. *β7* represents the impact of the AP STEM and English program on the trend (i.e., the slope) of AP exam taking/passing rates at each school during the three-year period after the program began.

Two parameters, *β4* and *β5*, play a role in establishing whether the treatment and control groups are balanced on both the level and trajectory of the outcome variable in the pre-intervention period. If these data were from a randomized controlled trial, we would expect similar levels and slopes prior to the intervention. However, in an observational study where equivalence between groups cannot be assumed, any observed differences will likely raise concerns about the ability to draw causal inferences about the relationship between the intervention and the outcomes (Linden and Adams, 2011).[[12]](#footnote-12) When the value for *β4* and/or *β5* is statistically significant, it indicates that, despite propensity score weighting, significant pre-intervention differences in AP STEM and English schools’ and comparison schools’ AP exam taking/passing rates remained.

# Appendix B

**Modeling Procedures for Difference-in-Difference (DID) Analyses**

DID is a quasi-experimental design that makes use of longitudinal data from treatment and control groups to estimate a causal effect. It calculates the effect of a treatment on an outcome by comparing the average change over time in the outcome variable for the treatment group, compared to the average change over time for the control group.

DID is used in observational settings where exchangeability cannot be assumed between the treatment and control groups. DID relies on a less strict exchangeability assumption (i.e., in absence of treatment, the unobserved differences between treatment and control groups are the same overtime). Hence, difference-in-difference is a useful technique to use when randomization at the level of the unit of measurement—in this case, the school level—is not possible. DID requires data from pre-/post-intervention, such as cohort or panel data or repeated cross-sectional data (individual or group level). The approach removes biases in post-intervention period comparisons between the treatment and control group that could be the result from permanent differences between those groups, as well as biases from comparisons over time in the treatment group that could be the result of trends due to other causes of the outcome. This method may still be subject to certain biases (e.g. mean regression, reverse causality and omitted variable bias), so results should be interpreted with caution.

**Modeling Procedure**

For each student group (all students, gender student groups, racial/ethnic student groups, student groups of interest, by subject) a DID model was developed to assess the impact of the AP STEM and English intervention on (a) school AP exam taking rates and (b) school AP exam passing rates at each school one year after the program began.

This procedure was used for all 24 of the DID analysis models developed for this report. The following equation represents the procedure:

*Yit = β0 + β1Timet + β2Participantit + β3TimetParticipantit + β4Covariatesi + eit*

In this model, Yit is the outcome measure for a school *i* at time *t*. *Timet* indicates if the measure occurred before or after the intervention began. *Participantit* is an indicator of whether or not a school was participating in the intervention at time *t*. *TimetParticipantit* is an interaction between *Timet* and *Participantit*. *Covariatesi* represents the vector of covariates included in the model for each school. Covariates considered for inclusion in the model included gender, race/ethnicity, English language learner status, and student disability status, with *eit* is the error term. In a DID model, the statistical significance of the interaction term (*TimetParticipantit*) is assessed to determine if there was a significant difference between treatment and control groups after treatment occurred.

The *β0* coefficient represents the baseline average in the outcome of interest. *β1* represents the change over time in the control group. *β2*represents the difference between the treatment and control groups prior to the intervention. *β3*represents the interaction between *Timet* and *Participantit* (i.e., the difference-in-differencebetween the treatment and control groups). *β4* represents the vector of covariates included in the model.

# Appendix C

**AP Exam Taking, Full Model Results**

**Table 3a: Impacts of AP STEM and English Program on AP Exam Taking, Any ELA, Math, or Science Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All∇** | **Female∇** | **Male∇** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White∇** | **ELL** | **Non-ELL** | **SWD⁰** | **Non-SWD∇** |
| Intercept (β0) | 3.91\*\*\* | 5.28\*\*\* | 2.28\*\* | 1.93\*\* | 10.62\*\*\* | 1.57\*\* | 2.28\* | 5.04\*\*\* | 0.55+ | 3.97\*\*\* | 0.11 | 4.13\*\*\* |
| (0.72) | (0.59) | (0.73) | (0.71) | (1.25) | (0.47) | (1.01) | (0.44) | (0.29) | (0.84) | (0.09) | (0.90) |
| Time (β1) | 1.14\*\*\* | 1.17\*\*\* | 0.98\*\* | 0.53\* | 1.40\*\* | 0.63\*\*\* | 0.60\* | 0.87\*\*\* | -0.01 | 1.21\*\* | 0.04 | 1.41\*\*\* |
| (0.33) | (0.21) | (0.34) | (0.21) | (0.54) | (0.16) | (0.29) | (0.10) | (0.11) | (0.39) | (0.03) | (0.40) |
| Intervention Period (β2) | -0.98+ | -0.88+ | -0.67 | -0.19 | -2.49+ | -1.04\* | 0.38 | -0.53\* | -0.54 | -0.92 | -0.02 | -1.27+ |
| (0.56) | (0.49) | (0.42) | (0.65) | (1.33) | (0.47) | (0.82) | (0.24) | (0.45) | (0.58) | (0.12) | (0.72) |
| Time by Intervention (β3) | -0.44 | -0.37 | -0.44 | -0.31 | -0.02 | -0.15 | 0.18 | -0.39\* | 0.19 | -0.33 | 0.07 | -0.52 |
| (0.28) | (0.24) | (0.33) | (0.41) | (0.70) | (0.22) | (0.51) | (0.19) | (0.25) | (0.35) | (0.06) | (0.34) |
| Participant (β4) | 0.48 | 0.19 | 0.90 | 0.72 | 0.43 | 0.39 | 1.77 | -0.15 | -0.17 | 0.51 | 0.08 | 0.96 |
| (0.83) | (0.77) | (0.81) | (1.00) | (2.73) | (0.74) | (1.65) | (0.68) | (0.49) | (0.94) | (0.15) | (1.01) |
| Participant by Time (β5) | -0.18 | -0.04 | -0.19 | 0.11 | -0.15 | 0.10 | 0.23 | 0.14 | 0.05 | -0.18 | -0.01 | -0.25 |
| (0.35) | (0.26) | (0.36) | (0.35) | (0.95) | (0.25) | (0.56) | (0.19) | (0.21) | (0.41) | (0.05) | (0.42) |
| Participation by Intervention (β6) | 6.39\*\*\* | 7.63\*\*\* | 4.82\*\*\* | 4.08\*\* | 10.97\*\*\* | 4.41\*\*\* | 2.54 | 6.13\*\*\* | 0.72 | 6.66\*\*\* | 0.46\* | 7.52\*\*\* |
| (0.73) | (0.81) | (0.59) | (1.36) | (2.66) | (1.08) | (1.71) | (0.60) | (0.70) | (0.74) | (0.18) | (0.91) |
| Participation by Time by Intervention (β7) | 0.52 | 0.37 | 0.66 | 0.55 | 0.11 | -0.13 | 2.02 | 0.71+ | 0.16 | 0.32 | 0.20 | 0.55 |
| (0.37) | (0.38) | (0.40) | (0.79) | (2.09) | (0.61) | (1.30) | (0.37) | (0.45) | (0.44) | (0.16) | (0.44) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | | | |

**Table 3b: Impacts of AP STEM and English Program on AP Exam Taking, Any ELA, Math, or Science Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Taking Change Two Years After Participation1** |
| All° | |  | | --- | | 0.06 | | (0.67) | | |  | | --- | | 0.78 | | (0.66) | | 4.35\*\*\*  (0.94) |
| Economically Disadvantaged | |  | | --- | | 0.20 | | (0.75) | | |  | | --- | | -1.45+ | | (0.75) | | 2.81\*\*  (1.06) |
| Non-Economically Disadvantaged | |  | | --- | | -0.02 | | (0.69) | | |  | | --- | | 1.18+ | | (0.68) | | |  | | --- | | 5.43\*\*\* | | (0.97) | |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1 Change in percentage points of students taking an AP exam at a school. A positive number indicates an increase in the percent taking for treatment schools. | | | |

**Table 4a: Impacts of AP STEM and English Program on AP Exam Taking, ELA Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All∇** | **Female∇** | **Male∇** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White∇** | **ELL⁰** | **Non-ELL** | **SWD⁰** | **Non-SWD∇** |
| Intercept (β0) | 1.17 | 2.32\*\*\* | 0.41 | 0.78\* | 3.96\*\*\* | 0.57 | 0.94 | 2.76\*\*\* | 0.02 | 1.60\*\* | -0.02 | 0.89 |
| (0.72) | (0.61) | (0.45) | (0.38) | (0.86) | (0.44) | (0.88) | (0.27) | (0.02) | (0.51) | (0.09) | (1.20) |
| Time (β1) | 1.13\*\*\* | 1.33\*\*\* | 0.76\*\*\* | 0.34\* | 0.79\* | 0.55\*\* | 0.21 | 0.77\*\*\* | -0.01 | 1.02\*\*\* | 0.06 | 1.51\*\* |
| (0.29) | (0.25) | (0.21) | (0.16) | (0.35) | (0.18) | (0.18) | (0.16) | (0.00) | (0.22) | (0.05) | (0.51) |
| Intervention Period (β2) | -1.59+ | -1.73\*\* | -0.89+ | 0.23 | -0.39 | -1.07\* | 0.32 | -0.62\*\* | 0.01 | -1.37\* | -0.16 | -2.47 |
| (0.86) | (0.65) | (0.51) | (0.32) | (0.85) | (0.45) | (0.58) | (0.22) | (0.01) | (0.56) | (0.15) | (1.51) |
| Time by Intervention (β3) | -0.73\*\* | -0.99\*\* | -0.46\* | -0.47 | -0.89 | -0.27 | -0.27 | -0.86\* | 0.01 | -0.51\* | 0.01 | -0.99\* |
| (0.27) | (0.37) | (0.20) | (0.38) | (0.57) | (0.27) | (0.32) | (0.39) | (0.00) | (0.23) | (0.06) | (0.46) |
| Participant (β4) | 0.63 | 0.31 | 0.59 | 0.53 | 0.11 | 0.28 | 1.87 | -0.36 | -0.41 | 0.23 | 0.02 | 1.23 |
| (0.79) | (0.75) | (0.50) | (0.74) | (1.61) | (0.66) | (1.38) | (0.53) | (0.32) | (0.61) | (0.12) | (1.25) |
| Participant by Time (β5) | -0.26 | -0.19 | -0.19 | 0.27 | 0.12 | 0.03 | 0.13 | 0.07 | 0.15 | -0.10 | 0.00 | -0.48 |
| (0.32) | (0.30) | (0.23) | (0.31) | (0.57) | (0.26) | (0.40) | (0.23) | (0.12) | (0.25) | (0.06) | (0.52) |
| Participation by Intervention (β6) | 5.15\*\*\* | 6.49\*\*\* | 3.35\*\*\* | 2.43\* | 8.87\*\*\* | 3.39\*\*\* | 2.92+ | 4.43\*\*\* | -0.12 | 5.15\*\*\* | 0.31+ | 6.63\*\*\* |
| (0.93) | (0.87) | (0.58) | (1.10) | (2.08) | (0.93) | (1.76) | (0.50) | (0.19) | (0.68) | (0.18) | (1.57) |
| Participation by Time by Intervention (β7) | 0.44 | 0.53 | 0.39 | 0.19 | -0.88 | 0.01 | 2.36\* | 0.80 | -0.11 | 0.14 | 0.07 | 0.59 |
| (0.36) | (0.49) | (0.29) | (0.67) | (1.28) | (0.55) | (0.94) | (0.49) | (0.18) | (0.34) | (0.11) | (0.54) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | | | |

**Table 4b: Impacts of AP STEM and English Program on AP Exam Taking, ELA Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Taking Change Two Years After Participation1** |
| All | 0.01  (0.50) | 0.00  (0.50) | 3.04\*\*\*  (0.71) |
| Economically Disadvantaged | 0.12  (0.53) | -1.03+  (0.53) | 2.46\*\*  (0.75) |
| Non-Economically Disadvantaged | -0.05  (0.54) | 0.10  (0.53) | 3.62\*\*\*  (0.76) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1 Change in percentage points of students taking an AP exam at a school. A positive number indicates an increase in the percent taking for treatment schools. | | | |

**Table 5a: Impacts of AP STEM and English Program on AP Exam Taking, Math Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All∇** | **Female∇** | **Male∇** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White⁰** | **ELL⁰** | **Non-ELL** | **SWD⁰** | **Non-SWD∇** |
| Intercept (β0) | 1.09\*\*\* | 1.30\*\*\* | 0.92\*\*\* | 2.97\*\* | 0.64\*\* | 0.33 | 1.22\*\*\* | 0.58+ | 1.20\*\*\* | 0.06 | 1.38\*\*\* |
| (0.16) | (0.16) | (0.18) | (1.06) | (0.21) | (0.54) | (0.26) | (0.32) | (0.19) | (0.05) | (0.19) |
| Time (β1) | 0.26\*\*\* | 0.26\*\*\* | 0.25\*\*\* | 0.52\*\* | 0.12+ | 0.16 | 0.24\*\*\* | -0.08 | 0.27\*\*\* | 0.01 | 0.31\*\*\* |
| (0.04) | (0.04) | (0.05) | (0.19) | (0.07) | (0.15) | (0.05) | (0.10) | (0.05) | (0.02) | (0.05) |
| Intervention Period (β2) | -0.07 | -0.04 | 0.06 | -0.60 | -0.03 | 0.48 | 0.02 | -0.17 | -0.04 | 0.04 | -0.05 |
| (0.10) | (0.12) | (0.16) | (0.52) | (0.21) | (0.50) | (0.13) | (0.43) | (0.13) | (0.05) | (0.12) |
| Time by Intervention (β3) | -0.02 | -0.02 | 0.01 | 0.15 | -0.08 | -0.27 | 0.00 | 0.09 | -0.02 | 0.02 | -0.01 |
| (0.07) | (0.08) | (0.08) | (0.37) | (0.11) | (0.30) | (0.10) | (0.07) | (0.08) | (0.04) | (0.09) |
| Participant (β4) | 0.49+ | 0.38 | 0.54\* | 2.45 | 0.10 | 1.63+ | 0.71+ | 0.02 | 0.43 | -0.01 | 0.50 |
| (0.26) | (0.31) | (0.27) | (1.66) | (0.42) | (0.92) | (0.39) | (0.39) | (0.29) | (0.08) | (0.31) |
| Participant by Time (β5) | 0.03 | 0.01 | 0.04 | -0.24 | 0.01 | 0.06 | 0.04 | 0.01 | 0.05 | 0.00 | 0.04 |
| (0.08) | (0.11) | (0.08) | (0.47) | (0.13) | (0.27) | (0.11) | (0.13) | (0.09) | (0.03) | (0.10) |
| Participation by Intervention (β6) | 2.17\*\*\* | 2.43\*\*\* | 1.85\*\*\* | 5.59\*\* | 1.25\*\* | 0.82 | 2.10\*\*\* | 0.21 | 2.28\*\*\* | 0.10 | 2.48\*\*\* |
| (0.32) | (0.42) | (0.31) | (1.77) | (0.45) | (0.92) | (0.39) | (0.49) | (0.34) | (0.10) | (0.38) |
| Participation by Time by Intervention (β7) | 0.33+ | 0.34 | 0.30 | 1.07 | 0.40 | 0.58 | 0.38 | 0.37+ | 0.29 | 0.10 | 0.36 |
| (0.19) | (0.24) | (0.21) | (1.51) | (0.32) | (0.75) | (0.24) | (0.23) | (0.20) | (0.12) | (0.22) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | | |

**Table 5b: Impacts of AP STEM and English Program on AP Exam Taking, Math Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Taking Change Two Years After Participation1** |
| All | 0.06  (0.39) | 0.86\*  (0.39) | 2.67\*\*\*  (0.55) |
| Economically Disadvantaged | 0.10  (0.38) | -0.18  (0.38) | 1.85\*\*\*  (0.54) |
| Non-Economically Disadvantaged | 0.03  (0.43) | 1.22\*\*  (0.42) | 2.96\*\*\*  (0.60) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students taking an AP exam at a school. A positive number indicates an increase in the percent taking for treatment schools. | | | |

**Table 6a: Impacts of AP STEM and English Program on AP Exam Taking, Science Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Female⁰** | **Male∇** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White∇** | **ELL** | **SWD⁰** | **Non-SWD⁰** |
| Intercept (β0) | 0.92\* | 0.78\*\*\* | 0.70\* | 4.10\* | 0.72\*\* | 1.04 | 1.38\*\*\* | 0.16 | 0.09+ | 0.95\*\* |
| (0.36) | (0.16) | (0.34) | (1.67) | (0.26) | (0.65) | (0.22) | (0.16) | (0.06) | (0.35) |
| Time (β1) | 0.17+ | 0.30\*\*\* | 0.14 | 0.53\* | 0.23\* | 0.14 | 0.31\*\*\* | 0.01 | 0.00 | 0.28\*\*\* |
| (0.09) | (0.06) | (0.11) | (0.26) | (0.10) | (0.22) | (0.05) | (0.07) | (0.02) | (0.07) |
| Intervention Period (β2) | 0.75 | -0.19 | 0.06 | -0.73 | -0.50 | -0.04 | -0.02 | -0.06 | 0.04 | -0.03 |
| (0.74) | (0.21) | (0.50) | (0.65) | (0.34) | (0.48) | (0.16) | (0.02) | (0.06) | (0.16) |
| Time by Intervention (β3) | -0.21 | 0.09 | 0.09 | 0.86 | 0.37\* | 0.34 | -0.03 | 0.04 | 0.07+ | 0.14 |
| (0.46) | (0.08) | (0.29) | (0.63) | (0.18) | (0.35) | (0.13) | (0.17) | (0.04) | (0.14) |
| Participant (β4) | 1.33\*\* | 0.53+ | 0.77 | 2.59 | 0.41 | 0.47 | 0.56 | 0.12 | 0.06 | 1.16\* |
| (0.48) | (0.27) | (0.58) | (2.77) | (0.43) | (0.92) | (0.36) | (0.22) | (0.09) | (0.46) |
| Participant by Time (β5) | 0.06 | 0.01 | -0.01 | -0.02 | 0.02 | 0.10 | 0.02 | -0.03 | -0.02 | 0.07 |
| (0.13) | (0.10) | (0.20) | (0.82) | (0.17) | (0.31) | (0.11) | (0.09) | (0.03) | (0.11) |
| Participation by Intervention (β6) | 2.19\*\* | 2.17\*\*\* | 1.96\* | 6.12\*\* | 1.72\* | 0.73 | 2.65\*\*\* | 0.28 | 0.15+ | 2.84\*\*\* |
| (0.84) | (0.37) | (0.84) | (2.06) | (0.67) | (1.01) | (0.47) | (0.42) | (0.09) | (0.41) |
| Participation by Time by Intervention (β7) | 0.85+ | 0.13 | 0.36 | -0.37 | -0.41 | 1.23 | 0.58\* | 0.20 | 0.03 | 0.33 |
| (0.50) | (0.19) | (0.53) | (1.60) | (0.40) | (0.79) | (0.25) | (0.33) | (0.07) | (0.24) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | |

**Table 6b: Impacts of AP STEM and English Program on AP Exam Taking, Science Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Taking Change Two Years After Participation1** |
| All° | -0.05  (0.43) | 0.73+  (0.43) | 0.43  (0.61) |
| Economically Disadvantaged | 0.04  (0.45) | -1.16\*  (0.45) | -0.08  (0.64) |
| Non-Economically Disadvantaged | -0.08  (0.46) | 0.80+  (0.46) | 1.27+  (0.65) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students taking an AP exam at a school. A positive number indicates an increase in the percent taking for treatment schools. | | | |

# Appendix D

**AP Exam Passing, Full Model Results**

**Table 7a: Impacts of AP STEM and English Program on AP Exam Passing, Any ELA, Math, or Science Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Female∇** | **Male∇** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White∇** | **ELL** | **Non-ELL** | **SWD** | **Non-SWD∇** |
| Intercept (β0) | 1.73\*\*\* | 0.74\* | 0.30 | 4.39\*\*\* | 0.26 | 1.12 | 1.85\*\*\* | 0.19+ | 1.32\*\* | 0.15 | 1.58\*\* |
| (0.26) | (0.34) | (0.19) | (0.86) | (0.21) | (0.78) | (0.27) | (0.10) | (0.39) | (0.10) | (0.55) |
| Time (β1) | 0.51\*\*\* | 0.45\*\* | 0.14\* | 1.03\*\* | 0.31\*\*\* | 0.29 | 0.40\*\*\* | -0.06 | 0.54\*\* | -0.03 | 0.70\*\* |
| (0.10) | (0.16) | (0.06) | (0.30) | (0.08) | (0.21) | (0.07) | (0.03) | (0.18) | (0.03) | (0.27) |
| Intervention Period (β2) | -0.05 | -0.12 | 0.28 | -0.52 | 0.03 | 1.24\* | -0.35\*\* | 0.13 | -0.10 | 0.15\* | -0.29 |
| (0.23) | (0.22) | (0.30) | (0.89) | (0.27) | (0.56) | (0.12) | (0.09) | (0.28) | (0.06) | (0.41) |
| Time by Intervention (β3) | -0.15 | -0.25 | -0.08 | -0.27 | -0.02 | 0.23 | -0.10 | 0.04 | -0.19 | 0.09+ | -0.29 |
| (0.12) | (0.21) | (0.12) | (0.54) | (0.14) | (0.44) | (0.07) | (0.06) | (0.16) | (0.04) | (0.22) |
| Participant (β4) | 0.66 | 0.78+ | 0.18 | 0.41 | 0.10 | 0.92 | 0.71+ | 0.10 | 0.85+ | -0.03 | 0.90 |
| (0.41) | (0.42) | (0.33) | (1.65) | (0.42) | (1.24) | (0.42) | (0.20) | (0.47) | (0.12) | (0.63) |
| Participant by Time (β5) | 0.03 | -0.03 | 0.05 | -0.03 | 0.08 | 0.18 | 0.14 | -0.01 | -0.03 | 0.01 | -0.11 |
| (0.14) | (0.18) | (0.11) | (0.58) | (0.17) | (0.41) | (0.14) | (0.07) | (0.19) | (0.04) | (0.28) |
| Participation by Intervention (β6) | 2.73\*\*\* | 1.92\*\*\* | 0.39 | 3.43\* | 1.09 | 0.31 | 2.94\*\*\* | 0.32 | 2.37\*\*\* | 0.03 | 2.84\*\*\* |
| (0.45) | (0.37) | (0.51) | (1.73) | (0.68) | (1.22) | (0.38) | (0.32) | (0.41) | (0.09) | (0.53) |
| Participation by Time by Intervention (β7) | 0.20 | 0.41 | 0.58+ | 0.93 | -0.26 | 0.42 | 0.23 | 0.00 | 0.29 | -0.01 | 0.39 |
| (0.21) | (0.27) | (0.32) | (1.46) | (0.42) | (0.85) | (0.22) | (0.25) | (0.21) | (0.07) | (0.27) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | | |

**Table 7b: Impacts of AP STEM and English Program on AP Exam Passing, Any ELA, Math, or Science Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Passing Change Two Years After Participation1** |
| All | 0.06  (0.41) | 0.61  (0.41) | 2.11\*\*\*  (0.59) |
| Economically Disadvantaged | 0.02  (0.38) | 0.05  (0.38) | 1.74\*\*  (0.54) |
| Non-Economically Disadvantaged | 0.03  (0.46) | 0.82+  (0.46) | 2.34\*\*\*  (0.66) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students passing an AP exam at a school. A positive number indicates an increase in the percent passing for treatment schools. | | | |

**Table 8a: Impacts of AP STEM and English Program on AP Exam Passing, ELA Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Female∇** | **Male∇** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White∇** | **Non-ELL** | **SWD** | **Non-SWD∇** | |
| Intercept (β0) | 1.28\*\*\* | 0.47\* | 0.12 | 1.67\*\* | 0.24 | 1.04+ | 1.48\*\*\* | 0.82\*\*\* | 0.06\* | 1.07\*\*\* | |
| (0.18) | (0.22) | (0.11) | (0.53) | (0.17) | (0.63) | (0.14) | (0.21) | (0.03) | (0.25) | |
| Time (β1) | 0.51\*\*\* | 0.33\*\* | 0.06 | 0.72\*\* | 0.19\*\* | 0.08 | 0.40\*\*\* | 0.43\*\*\* | -0.01+ | 0.52\*\*\* | |
| (0.08) | (0.12) | (0.05) | (0.22) | (0.06) | (0.12) | (0.06) | (0.10) | (0.01) | (0.12) | |
| Intervention Period (β2) | -0.33+ | -0.41 | 0.26 | -0.18 | -0.12 | 0.60 | -0.39\*\* | -0.32 | 0.07\*\* | -0.45 | |
| (0.19) | (0.35) | (0.17) | (0.61) | (0.23) | (0.46) | (0.12) | (0.29) | (0.03) | (0.33) | |
| Time by Intervention (β3) | -0.27\* | -0.22 | -0.07 | -0.78\* | -0.01 | -0.15 | -0.32\*\* | -0.19\* | 0.03 | -0.24\* | |
| (0.11) | (0.14) | (0.10) | (0.39) | (0.11) | (0.32) | (0.09) | (0.09) | (0.02) | (0.10) | |
| Participant (β4) | 0.17 | 0.18 | 0.15 | -0.06 | 0.23 | 0.90 | -0.11 | 0.25 | 0.00 | 0.22 | |
| (0.34) | (0.26) | (0.27) | (1.12) | (0.37) | (0.99) | (0.31) | (0.29) | (0.05) | (0.34) | |
| Participant by Time (β5) | 0.03 | -0.04 | 0.05 | -0.03 | 0.00 | 0.02 | 0.10 | 0.01 | 0.00 | -0.01 | |
| (0.13) | (0.13) | (0.11) | (0.45) | (0.14) | (0.29) | (0.12) | (0.13) | (0.02) | (0.14) | |
| Participation by Intervention (β6) | 2.23\*\*\* | 1.60\*\*\* | 0.42 | 4.42\*\* | 0.92 | 1.44 | 2.06\*\*\* | 1.91\*\*\* | 0.05 | 2.21\*\*\* | |
| (0.40) | (0.40) | (0.44) | (1.60) | (0.59) | (1.09) | (0.32) | (0.37) | (0.05) | (0.43) | |
| Participation by Time by Intervention (β7) | -0.01 | 0.02 | 0.31 | -0.60 | -0.11 | 0.88 | 0.12 | -0.07 | 0.01 | -0.08 | |
| (0.20) | (0.19) | (0.26) | (1.00) | (0.36) | (0.71) | (0.20) | (0.16) | (0.05) | (0.18) | |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | |

**Table 8b: Impacts of AP STEM and English Program on AP Exam Passing, ELA Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Passing Change Two Years After Participation1** |
| All | -0.01  (0.28) | 0.14  (0.28) | 1.24\*\*  (0.40) |
| Economically Disadvantaged | -0.01  (0.28) | -0.10  (0.28) | 1.39\*\*\*  (0.40) |
| Non-Economically Disadvantaged | -0.04  (0.32) | 0.24  (0.32) | 1.19\*\*  (0.45) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students passing an AP exam at a school. A positive number indicates an increase in the percent passing for treatment schools. | | | |

**Table 9a: Impacts of AP STEM and English Program on AP Exam Passing, Math Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All°** | **Female∇** | **Male⁰** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White** | **ELL** | **Non-ELL** | **SWD** | **Non-SWD⁰** |
| Intercept (β0) | -0.17+ | 0.41\*\*\* | -0.19+ | 0.00 | 1.89\*\* | 0.17+ | 0.33 | 0.76\*\*\* | 0.19 | 0.48\*\*\* | 0.04+ | -0.18 |
| (0.10) | (0.09) | (0.11) | (0.08) | (0.66) | (0.10) | (0.36) | (0.10) | (0.12) | (0.10) | (0.02) | (0.13) |
| Time (β1) | 0.10\*\* | 0.08\*\* | 0.12\*\*\* | 0.08\* | 0.27+ | 0.08\* | 0.02 | 0.14\*\*\* | -0.06 | 0.13\*\*\* | -0.01 | 0.12\*\*\* |
| (0.03) | (0.03) | (0.04) | (0.03) | (0.15) | (0.04) | (0.07) | (0.03) | (0.04) | (0.03) | (0.01) | (0.03) |
| Intervention Period (β2) | 0.10 | 0.15+ | 0.08 | 0.18 | 0.06 | 0.09 | 0.94\* | 0.06 | 0.15 | 0.10 | 0.05\* | 0.11 |
| (0.07) | (0.08) | (0.09) | (0.20) | (0.36) | (0.16) | (0.39) | (0.08) | (0.14) | (0.07) | (0.02) | (0.08) |
| Time by Intervention (β3) | 0.05 | 0.07 | 0.02 | -0.12 | 0.29 | -0.05 | 0.01 | 0.01 | 0.01 | 0.03 | 0.03\* | 0.04 |
| (0.05) | (0.05) | (0.06) | (0.10) | (0.28) | (0.07) | (0.26) | (0.05) | (0.04) | (0.05) | (0.02) | (0.05) |
| Participant (β4) | 0.14 | 0.14 | 0.24 | 0.02 | 0.95 | 0.00 | 0.61 | 0.14 | 0.04 | 0.20 | 0.01 | 0.20 |
| (0.17) | (0.17) | (0.19) | (0.11) | (1.10) | (0.22) | (0.76) | (0.19) | (0.21) | (0.16) | (0.04) | (0.20) |
| Participant by Time (β5) | 0.03 | 0.01 | 0.03 | 0.02 | -0.04 | 0.02 | 0.08 | 0.03 | 0.00 | 0.01 | 0.00 | 0.03 |
| (0.04) | (0.05) | (0.05) | (0.06) | (0.34) | (0.08) | (0.17) | (0.05) | (0.08) | (0.04) | (0.01) | (0.05) |
| Participation by Intervention (β6) | 0.72\*\*\* | 0.89\*\*\* | 0.56\*\* | -0.11 | 1.65 | 0.10 | -0.63 | 0.90\*\*\* | -0.06 | 0.73\*\*\* | -0.03 | 0.83\*\*\* |
| (0.15) | (0.20) | (0.17) | (0.27) | (1.34) | (0.26) | (0.63) | (0.18) | (0.17) | (0.15) | (0.04) | (0.18) |
| Participation by Time by Intervention (β7) | 0.08 | -0.08 | 0.25\* | 0.18 | 0.36 | 0.13 | 0.30 | 0.05 | 0.22 | 0.08 | 0.00 | 0.11 |
| (0.10) | (0.11) | (0.13) | (0.24) | (1.09) | (0.19) | (0.47) | (0.12) | (0.16) | (0.10) | (0.03) | (0.11) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | | | |

**Table 9b: Impacts of AP STEM and English Program on AP Exam Passing, Math Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Passing Change Two Years After Participation1** |
| All | 0.02  (0.23) | 0.47\*  (0.22) | 1.21\*\*\*  (0.32) |
| Economically Disadvantaged | 0.01  (0.20) | -0.07  (0.20) | 0.75\*\*  (0.28) |
| Non-Economically Disadvantaged | 0.01  (0.26) | 0.65\*  (0.26) | 1.38\*\*\*  (0.37) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students passing an AP exam at a school. A positive number indicates an increase in the percent passing for treatment schools. | | | |

**Table 10a: Impacts of AP STEM and English Program on AP Exam Passing, Science Exams—CITS**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All⁰** | **Female⁰** | **Male⁰** | **Afr. Amer. / Black⁰** | **Asian⁰** | **Hispanic/ Latino** | **Multi-Race⁰** | **White⁰** | **ELL** | **Non-ELL** | **SWD** | **Non-SWD⁰** |
| Intercept (β0) | -0.27\*\* | -0.09 | -0.31\*\*\* | -0.01 | 1.52 | 0.05 | 0.54 | 0.23 | 0.07 | 0.31\*\* | 0.02 | -0.25\* |
| (0.09) | (0.13) | (0.08) | (0.06) | (0.97) | (0.11) | (0.51) | (0.23) | (0.06) | (0.12) | (0.01) | (0.12) |
| Time (β1) | 0.09\*\* | 0.07\* | 0.12\*\*\* | 0.09\*\* | 0.00\*\*\* | 0.11\* | 0.02 | 0.12\*\*\* | -0.02 | 0.11\*\* | 0.00 | 0.11\*\* |
| (0.03) | (0.04) | (0.03) | (0.03) | (0.00) | (0.04) | (0.17) | (0.04) | (0.02) | (0.03) | (0.00) | (0.04) |
| Intervention Period (β2) | 0.09 | 0.12 | 0.10 | -0.09 | 0.43\*\* | 0.00 | 0.54 | 0.13 | 0.05 | 0.10 | 0.06 | 0.12 |
| (0.09) | (0.10) | (0.10) | (0.12) | (0.16) | (0.14) | (0.36) | (0.09) | (0.05) | (0.10) | (0.04) | (0.11) |
| Time by Intervention (β3) | 0.09+ | 0.14\* | 0.02 | -0.10+ | -0.07 | 0.12 | 0.22 | -0.04 | 0.02 | 0.10+ | 0.03 | 0.10+ |
| (0.05) | (0.06) | (0.05) | (0.06) | (0.48) | (0.09) | (0.30) | (0.07) | (0.03) | (0.05) | (0.03) | (0.06) |
| Participant (β4) | 0.07 | 0.15 | 0.16 | 0.04 | 0.06 | 0.05 | 0.07 | 0.72\* | 0.02 | 0.34+ | 0.01 | 0.15 |
| (0.16) | (0.20) | (0.16) | (0.24) | (0.37) | (0.24) | (0.68) | (0.28) | (0.09) | (0.17) | (0.04) | (0.20) |
| Participant by Time (β5) | 0.00 | -0.01 | 0.01 | 0.02 | 0.48 | 0.02 | 0.10 | -0.03 | 0.00 | -0.01 | 0.00 | 0.01 |
| (0.05) | (0.06) | (0.05) | (0.11) | (1.18) | (0.09) | (0.24) | (0.07) | (0.03) | (0.05) | (0.01) | (0.06) |
| Participation by Intervention (β6) | 0.93\*\*\* | 0.93\*\*\* | 0.90\*\*\* | 0.17 | 0.18 | 0.46 | -0.06 | 1.12\*\*\* | 0.23 | 0.94\*\*\* | 0.00 | 1.06\*\*\* |
| (0.20) | (0.23) | (0.22) | (0.47) | (0.40) | (0.37) | (0.88) | (0.27) | (0.30) | (0.20) | (0.05) | (0.23) |
| Participation by Time by Intervention (β7) | 0.17+ | 0.25+ | 0.07 | 0.38+ | -0.13 | -0.13 | 0.21 | 0.27+ | -0.13 | 0.16 | -0.02 | 0.19 |
| (0.10) | (0.13) | (0.12) | (0.21) | (1.15) | (0.24) | (0.53) | (0.14) | (0.19) | (0.10) | (0.05) | (0.12) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  °After propensity score weighting, treatment and comparison schools were only partially balanced.  ∇Alternate weighting was used. See methods for description.  β6 represents the difference in the level between AP STEM and English schools and comparison schools during the year immediately following intervention (i.e. SY12 for Cohort III). β7 represents the difference between AP STEM and English schools and comparison schools in the slope from the years before the AP STEM and English intervention (i.e., SY07–SY10 for Cohort III) to the three years after the AP STEM and English intervention began (i.e., SY12–SY13 for Cohort III). | | | | | | | | | | | | |

**Table 10b: Impacts of AP STEM and English Program on AP Exam Passing, Science Exams—DID (Cohorts VIII and IX)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description**  **(Student group)** | **Participant** | **Time** | **Percent Passing Change Two Years After Participation1** |
| All | 0.02  (0.22) | 0.41+  (0.22) | 0.20  (0.31) |
| Economically Disadvantaged | 0.01  (0.17) | 0.16  (0.17) | 0.00  (0.24) |
| Non-Economically Disadvantaged | -0.01  (0.25) | 0.47+  (0.25) | 0.39  (0.36) |
| + *p* < .1, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001  1Change in percentage points of students passing an AP exam at a school. A positive number indicates an increase in the percent passing for treatment schools. | | | |

# Appendix E

**Portion of schools experiencing increased exam participation**

**Table 11: Portion of Participating Schools Experiencing Increased AP Exam Participation During First Year of Participation, by Number of Subject Areas**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **School Year** | **No Subject Areas** | **1 AP Subject Area** | **2 AP Subject Areas** | **3 AP Subject Areas** |
| SY09 | 0% | 13% | 25% | 63% |
| SY10 | 0% | 0% | 0% | 100% |
| SY11 | 0% | 4% | 35% | 62% |
| SY12 | 0% | 0% | 22% | 78% |
| SY13 | 0% | 0% | 23% | 77% |
| SY14 | 0% | 8% | 15% | 77% |
| SY15 | 0% | 6% | 33% | 61% |
| SY16 | 0% | 0% | 33% | 67% |
| SY17 | 30% | 20% | 30% | 20% |
| SY18 | 33% | 17% | 33% | 17% |
| SY19 | 0% | 40% | 60% | 0% |
| SY09–SY19 | 4% | 7% | 27% | 62% |

**Table 12: Portion of Participating Schools Experiencing Increased AP Exam Participation During First Year of Participation, by AP Exam Subject**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **School Year** | **ELA, Math, or Science** | **ELA** | **Math** | **Science** |
| SY09 | 100% | 100% | 75% | 75% |
| SY10 | 100% | 100% | 100% | 100% |
| SY11 | 96% | 92% | 85% | 81% |
| SY12 | 89% | 78% | 100% | 100% |
| SY13 | 92% | 92% | 85% | 100% |
| SY14 | 100% | 100% | 92% | 77% |
| SY15 | 89% | 78% | 94% | 89% |
| SY16 | 100% | 100% | 100% | 67% |
| SY17 | 60% | 70% | 50% | 20% |
| SY18 | 50% | 33% | 67% | 33% |
| SY19 | 80% | 100% | 20% | 40% |
| SY09–SY19 | 90% | 88% | 84% | 77% |

1. Increasing participation and performance in English Laguage Arts (ELA) AP courses and AP exams is not a stated goal of the project. However, increasing participation and performance in ELA AP courses and exams is frequently reflected as a goal of the program in practice. [↑](#footnote-ref-1)
2. Data from the 2019–20 school year were not included because the necessary data was not available. [↑](#footnote-ref-2)
3. Ten participating schools from Cohort III through Cohort IX were not included in the treatment group because they were identified as not having fully implemented the program by DESE and Mi. These schools included Frontier Regional High School, Hoosac Valley Middle and High School, KIPP Academy Lynn Collegiate High School, Milford High School, Rockland Senior High School, Stoughton High School, Wachusett Regional High School, Weymouth High School, and William J. Dean Technical High School. [↑](#footnote-ref-3)
4. To be considered an exam taker, a student must have taken one or more AP ELA, math, or science exam. To be considered an exam passer, a student must have scored a three or better on at least one of their AP ELA, math, or science exams. [↑](#footnote-ref-4)
5. AP exam passing rates are calculated based on the number of students enrolled in the school rather than the number of students taking the AP exam, since the objective of the analysis was to determine if the *overall rate of AP exam passing* had changed more at participating schools than at similar non-participating schools, and not to determine if the *rate of passing among exam takers* had changed more at participating schools than at similar non-participating schools. [↑](#footnote-ref-5)
6. CITS analyses for other student groups—such as non-binary students—were not conducted because the amount of data necessary to conduct the analysis was not available. [↑](#footnote-ref-6)
7. It is important to note that propensity score *weighting* is a separate and different procedure than propensity score *matching*. While there are advantages and disadvantages of both procedures, ultimately, propensity score weighting was used because it resulted in low standardized bias and a high number of balanced models. [↑](#footnote-ref-7)
8. Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3–4), 169–188. [↑](#footnote-ref-8)
9. Corresponding to *β4* and *β5* coefficients for the CITS analysis, see Appendix A, and *β*2 coefficient for the DID analysis, see Appendix B. [↑](#footnote-ref-9)
10. DESE began using ECODIS as a proxy for low-income in SY15. [↑](#footnote-ref-10)
11. Interpreting signifance of results that relied on alternate weighting should be done with caution as the surety of the comparability between the treatment and comparison groups is not guaranteed. [↑](#footnote-ref-11)
12. Linden, A., & Adams, J. L. (2011). Applying a propensity score‐based weighting model to interrupted time series data: Improving causal inference in programme evaluation. *Journal of Evaluation in Clinical Practice*, *17*(6), 1231–1238. [↑](#footnote-ref-12)