



**Comparative Evaluation:
Participants Versus Nonparticipants in the Achieving
Collegiate Excellence and Success (ACES) Program at
Montgomery County Public Schools in
Year One and Year Two**

Office of Shared Accountability

April 2016

**Shahpar Modarresi, Ph.D.
Natalie Wolanin
Elizabeth Cooper-Martin, Ph.D.**



OFFICE OF SHARED ACCOUNTABILITY

850 Hungerford Drive
Rockville, Maryland 20850
301-279-3553

Mr. Larry A. Bowers
Interim Superintendent of Schools

Dr. Maria V. Navarro
Chief Academic Officer

Table of Contents

Executive Summary	v
Evaluation Design and Method.....	v
Summary of Findings.....	vi
Conclusion	vi
Program Description	1
Literature Highlights	2
Evaluation Questions	4
Study Design and Methods	4
Design	4
Data Sources	5
Study Samples.....	5
Outcome Measures.....	5
Analytical Procedures	6
Question 1	6
Question 2	7
Strengths and Limitations	8
Organization of Findings	9
Findings for Year Two.....	10
Question 1. Are there significant differences in high school outcomes between ACES and a comparison group of students?.....	10
Grade 11	10
Grade 12.....	12
Question 2. Are there significant differences in college-readiness outcomes between ACES and a comparison groups of students?.....	13
Grade 11	13
Grade 12.....	14
Findings for Year One	16
Question 1: Are there significant differences in high school outcomes between ACES students and a comparison group of students?.....	16
Grade 11	16
Grade 12.....	17
Question 2: Are there significant differences in college-readiness outcomes between ACES students and a comparison group of students?	19
Grades 11 and 12	19

Questions 1 and 2.....20

Summary of Findings..... 21

Study Conclusion 21

Acknowledgements..... 22

References..... 23

Appendix A: Characteristics of ACES and Comparison Students 26

Appendix B: Methodology and Findings for School-level Matching Comparison Group in Year One..... 28

 Methodology28

 Design 28

 Study Sample 29

 Analytical Procedures 29

Findings for Students in the ACES Program vs. Students in the School-level Matched Group31

 Question 1. Are there significant differences in high school outcomes between ACES and a comparison group of students? 31

 Question 2. Are there significant differences in college-readiness outcomes between ACES and a comparison group of students?..... 34

List of Tables

Table 1 ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores Between ACES and a Comparison Group of Grade 11 Students in Year Two	11
Table 2 ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores Between ACES and a Comparison Group of Grade 12 Students in Year Two	12
Table 3 Logistic Regression Results Comparing Promotion, SAT or ACT Participation, SAT/ACT Milestone, and AP/IB Milestone Between ACES and a Comparison Group of Grade 11 Students for Year Two.....	14
Table 4 Logistic Regression Results Comparing Promotion, SAT and ACT Participation, SAT and ACT Milestone, and AP or IB Milestone Between ACES and a Comparison Group of Grade 12 Students in Year Two.....	15
Table 5 ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores Between ACES and a Comparison Group for Grade 11 in Year One.....	16
Table 6 ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores Between ACES and a Comparison Group for Grade 12 in Year One.....	18
Table 7 Promotion to Next Level Between ACES and a Comparison Group by Grade Level in Year One.....	19
Table 8 Participation in SAT or ACT Between ACES and a Comparison Group by Grade Level in Year One.....	20
Table 9 Met SAT/ACT Milestone Between ACES and a Comparison Group of Test Takers by Grade Level in Year One.....	20
Table 10 Met AP/IB Milestone Between ACES and a Comparison Group by Grade Level in Year One.....	20
Table A1 Characteristics of All ACES Students, Matched ACES Students, and Individual-level Matched Comparison Students for Year Two.....	26
Table A2 Characteristics of All ACES Students, Matched ACES Students, and Individual-level Matched Comparison Students for Year One.....	27
Table B1 Characteristics of All ACES Students, Matched ACES Students, and School-level Matched Comparison Students for Year One.....	29
Table B2 Comparisons of GPA, Quarter 4 MPA, Unexcused Absences, SAT Scores, and ACT Scores Between ACES and School-level Matched Group for Grade 11 in Year One.....	31
Table B3 Comparisons of GPA, Quarter 4 MPA, Unexcused Absences, SAT Scores, and ACT Scores Between ACES and Matched Groups of Students for Grade 12 in Year One.....	33

Table B4 Promotion to Next Level Between ACES and a Comparison Group by Grade Level in Year One..... 34

Table B5 Participation in SAT or ACT Between ACES and a Comparison Group by Grade Level in Year One..... 35

Table B6 Met SAT/ACT Milestone Between ACES and a Comparison Group of Test Takers by Grade Level in Year One 35

Table B7 Met AP/IB Milestone Between ACES and a Comparison Group by Grade Level in Year One..... 35

Executive Summary

The Office of Shared Accountability in Montgomery County Public Schools (MCPS) has been conducting a multiyear evaluation of the Achieving Collegiate Excellence and Success (ACES) program. ACES is a program that was instituted in collaboration with Montgomery College (MC) and the Universities at Shady Grove (USG) to support students from backgrounds that are underrepresented in higher education. Year One of the program was 2013–2014 and Year Two was 2014–2015. This study, one in a series of evaluations of the ACES program, addresses the following questions about the immediate outcomes of the ACES program for students in MCPS: 1) Are there significant differences in high school outcomes between ACES and a comparison group of students? 2) Are there significant differences in college-readiness outcomes between ACES and a comparison group of students?

Evaluation Design and Method

Design

Because students were not randomly assigned to the ACES program, this evaluation used a quasi-experimental design. The key component of such a design is appropriate comparison groups when evaluating a program's outcomes. Therefore, comparison groups were created using multivariate statistical techniques. The outcomes of ACES students were compared with those of their peers in an individual-level matched group of students for Year One and also for Year Two. Additionally, the outcomes of ACES students in Year One were compared to a second group, a school-level matched group of students.

Measures

The measures of immediate high school outcomes were end-of-year grade point average (GPA), Quarter 4 marking period average (Q4 MPA), sum of unexcused absences, and scores on ACT and SAT tests. The measures of immediate college-readiness outcomes were as follows: promotion to next level (i.e., promotion to 12th grade for 11th graders, graduation for 12th graders); completing SAT or ACT tests; attainment of college-readiness milestones for SAT/ACT, for Advanced Placement or International Baccalaureate (AP/IB) exams, and for ACCUPLACER placement tests.

Participating Schools and Students

Students in Grades 11 and 12 participated in Year One and Year Two of the ACES program at 10 high schools: Montgomery Blair, Clarksburg, Albert Einstein, Gaithersburg, John F. Kennedy, Northwood, Rockville, Seneca Valley, Wheaton, and Watkins Mill. Across all 10 schools, 968 students participated in Year One and 1,019 in Year Two. Because matches were not found for each ACES student, the study samples were smaller: the Year One sample included 793 ACES and 793 comparison students and the Year Two sample included 779 ACES and 779 comparison students.

Analytical Procedures

Analyses were conducted using multivariate (analysis of covariance and logistic regression) and descriptive statistics, as appropriate. Significance tests from the multivariate procedures were

supplemented by reports of effect sizes. Effect sizes were used to judge whether differences between the groups of students were large enough to be of practical significance to educators.

Summary of Findings

A summary of findings using individual-level comparison groups follows. Findings using school-level matched comparison groups were somewhat similar, as detailed in Appendix B.

Question 1. Are there significant differences in high school outcomes between ACES and a comparison group of students?

The analyses on high school outcomes revealed no significant differences between ACES students and comparison students in both Year One and Year Two on two measures: end-of-year GPA and Q4 MPA. However, there were significant differences between ACES students and their peers on the sum of unexcused absences; 11th graders in the ACES group had fewer unexcused absences in Year Two, and 12th graders in the ACES group had fewer unexcused absences in both Year One and Year Two. Lastly, there were significant differences in favor of the comparison group on SAT and ACT scores in Year One for both Grades 11 and 12; but in Year Two, the only significant difference was on ACT scores for 11th graders, again in favor of comparison students. Further, the effect sizes from the analyses of these SAT and ACT scores decreased from Year One to Year Two for both grade levels, indicating that the differences between the ACES and comparison students narrowed.

Question 2. Are there significant differences in college-readiness outcomes between ACES and a comparison group of students?

Analysis of college-readiness outcomes revealed significant differences in promotion to the next level as follows: in Year One, ACES students in Grade 11 were more likely to be promoted, and in year Two, ACES students in Grade 12 were more likely to graduate. In both years, there were significant differences in participation in the SAT or ACT; more ACES than comparison students in both Grade 11 and Grade 12 took the SAT or ACT. Among students who took the SAT/ACT, the percentage of ACES students who met the MCPS milestone was significantly lower than the comparison group for both Grades 11 and Grade 12 in Year One. But in Year Two, there were no significant differences between ACES and comparison students in meeting the SAT/ACT milestone. Similarly, there were no significant differences between ACES and comparison students in meeting the AP/IB milestone in either year or in meeting the ACCUPLACER benchmarks (analyzed only for Year Two).

Conclusion

In conclusion, this report does not provide evidence that the ACES program improved students' grades or scores on college-readiness tests (i.e., ACT, SAT, ACCUPLACER). However, it is worth noting that these measures reflect students' decisions, achievement, coursework, and other experiences prior to their entry into the ACES program and that, during Year One, the 12th graders in ACES participated in the program for only one year. It is a positive sign that the differences

between the ACES students and their peers on ACT and SAT scores decreased in the second year of program implementation.

Further, there is evidence that ACES positively impacted students' decisions while in the program. For example, it appears that ACES students were more conscientious about school attendance (i.e., had fewer unexcused absences) than their peers. Further, ACES students were more likely than their non-ACES peers to take the steps necessary for completing the ACT or SAT. This finding suggests that ACES students were more intent on attending college. Additionally, as presented in a companion report (Wolanin & Cooper-Martin, 2016), almost all of the 12th grade ACES students in Year Two applied to, gained admission to, and reported plans to attend college. Unfortunately, we were unable to test whether ACES students were more likely than their peers to accomplish these steps, due to a lack of information about comparison students.

Finally, there was evidence for a positive impact of ACES on promotion and graduation. The strongest evidence related to high school graduation, an accomplishment that reflected students' behavior and decisions while in the ACES program, as well as in prior years, and which is essential for reaching the program's long-term goal that each ACES student earn a 4-year college degree.

Comparative Evaluation: Participants Versus Nonparticipants in the Achieving Collegiate Excellence and Success (ACES) Program at Montgomery County Public Schools in Year One and Year Two

Achieving Collegiate Excellence and Success (ACES) is a collaborative effort among Montgomery County Public Schools (MCPS), Montgomery College (MC), and the Universities at Shady Grove (USG). Using a case management approach, the ACES program seeks to create a seamless pathway from high school to college completion. This free program targets students in Grades 11 and 12 and identifies and supports both students who come from backgrounds that are underrepresented in higher education and those who would be the first in their family to attend college. According to a Memorandum of Understanding (MOU) between MCPS, MC, and USG, MCPS will lead the design of research protocols in consultation with all partners to evaluate the effectiveness of the program (MCPS, MC, & USG, 2013). Consequently, the Office of Shared Accountability in MCPS is conducting a multiyear evaluation of the ACES program. This report is one in a series of studies which present the findings associated with the immediate outcomes of the ACES program in Year One (2013–2014) and Year Two (2014–2015) of implementation at MCPS. The more specific objective of this report is to compare the immediate outcomes of ACES students to comparison groups of students; individual-level matched comparison students were used in Year One and in Year Two, and school-level matched comparison students were used in Year One (located in the appendix).

Program Description

ACES is a student support program that provides direct academic and student support to eligible 11th and 12th graders in 10 MCPS high schools. The broad goal of ACES is to identify and neutralize barriers that may prevent at-risk students from graduating with bachelor degrees. A central element to the ACES program is the presence of MC coaches who mentor, advocate, and advise ACES students regarding college and college-readiness activities. The coaches provide support using a case management approach where coaches work individually with students and provide as-needed support to ensure success in high school, complete college admission applications, negotiate the scholarship and financial aid process, and transition from high school to college, as well as provide weekly activities, group meetings and college trips. Other activities include offering summer programs at MC and USG and providing transitional services to college. These ACES program activities are in addition to and align with the college and career supports provided by MCPS staff members. Students are expected to commit to their own academic success and participate in planned activities as well as meet with their ACES coach on a regular basis.

The ACES program began in the fall of 2013 at 10 high schools. Students had to apply to the program, and most did so in the spring of their sophomore year. If accepted, they were assigned to the ACES coach at their school during their junior and senior years. In Year Two, a new cohort of Grade 11 students began the ACES program. For those ACES students who choose to attend MC and USG, the ACES program continues to provide post-secondary support and coaching.

Participating Schools and Students

The following 10 MCPS high schools participated in both years of the ACES program: Montgomery Blair, Clarksburg, Albert Einstein, Gaithersburg, John F. Kennedy, Northwood, Rockville, Seneca Valley, Wheaton, and Watkins Mill. It was expected that 60 students from Grade 11 and 60 students from Grade 12 at each school would participate in ACES for an approximate total of 1,200 students. Specifically, a total of 968 11th and 12th graders in these schools participated in the first year (2013–2014) of the ACES program and 1,019 in the second year (2014–2015). Each ACES student met one or more of the following risk factors: member of an underrepresented race/ethnicity group in higher education; low income or single parent household; first generation college student; student receiving special education services¹; immigrant or child of immigrant parents; and homeless or living in unstable conditions.

Literature Highlights

It is imperative for the economic and social well-being of the nation that college access be increased for all students, but more steeply for underrepresented minorities placed at risk. Recently, there has been an influx of programs that aimed to increase college readiness and access for students. However, in May of 2012, Thomas Bailey, director of the National Center for Postsecondary Research (NCPR), noted the scarcity of empirical work evaluating the effects of college readiness partnerships/programs (Barnett, 2012). Nevertheless, he states that there is reason to believe that these programs do indeed positively impact college readiness. Empirical studies will ensure that program elements that are shown to be instrumental elsewhere are replicated, and those that are not are altered or eradicated.

In a study by Cates and Schaeffle (2011), a low-income, 70% Latino student population participated in a college preparation program for six years. Components of this program included, but were not limited to, tutoring, mentoring, advising and summer programs. The goal of the study was four fold in that it aimed to examine: 1) the relationship between components of the program and college track course completion, 2) the relationship between hours participated in components of the program and early PSAT exposure, 3) the elements of the program that lead to college success, and 4) the relationship between program elements and self-report expectations of students going to college. Approximately 187 students that attended school in one of four school districts in a rural, western region of the U.S. participated in the program. At the onset of data collection, students were in the 5th and 6th grades. Data were collected on students over the six year period of participation in the program. Results of the study found that hours of advising were positively correlated to completing college-track classes, whereas tutoring was negatively related to the completion of college-track classes. The researchers suggest that the finding implies the value of advising for minority students in college preparation and the possibility that students who struggle academically and need more tutoring are less likely to enter college-track classes. The study also found that students who were less likely to participate in the program's elements were less likely to take the PSAT. The researchers also noted that students who participated in the program

¹ Special education services provide specially designed instruction that involves modifications to the curriculum itself, to the way the curriculum is taught, or both, in order to meet the specific needs of the student. Students may also receive other special education-related services (e.g., speech and language therapy or occupational therapy) (MCPS, 2015).

reported that visiting college campuses, listening to speakers from other colleges, and receiving pamphlets about college preparation were most influential in their decision to go to college.

The above study highlighted the importance of early preparation for college readiness in Latino and low-income students. However, it is important to note that the study reflected students' trajectory from 5th grade through high school.

In an empirical study that used a nationally representative sample of high school seniors, Engbert and Wolniak (2009) found that race did not influence college enrollment, although race groups differed in destinations, and that socio-economic status was associated with college enrollment. Additionally, college affordability increased enrollment into any college, but decreased the odds of attending a 4-year college. They found the greatest impact on college enrollment was academic achievement, college aspirations of family and friends, and college-linking activities. College-linking activities included networking with teachers, counselors, peers, parents, and college representatives. The authors state that "our findings underscore the importance of motivating students early on to seek out college information and to discuss their school and college plans with family members" (Engbert & Wolniak, p. 148). Their study also found that certain high school variables influenced college enrollment, such as the socioeconomic level of the school's student body and a culture where students' families and friends aspire for college enrollment.

For more review of the relevant literature, please see Wolanin and Modarresi (2015).

Evaluation Questions

The current study addressed the following questions about the immediate outcomes of the ACES program in its first two years of implementation at MCPS:

Question 1. Are there significant differences in high school outcomes between ACES and a comparison group of students?

Question 2. Are there significant differences in college-readiness outcomes between ACES and a comparison group of students?

Study Design and Methods

Design

The critical issue in outcome evaluations is whether or not a program produces levels of effects above what would have occurred without the program (Rossi & Freeman, 1993). Many researchers assert that randomized studies via experimental designs are necessary to reliably determine program outcomes or effects (Ashenfelter & Card, 1985; Barnow, 1987; Burtless & Orr, 1986; Hedrick, Bickman, & Rog, 1993; Shadish, Cook, & Campbell, 2002). These researchers argue that only random assignment of students either to a treatment (i.e., ACES program) group or to a comparison (i.e., regular school activities) group will control for all confounding variables or guard against each of the sources of internal invalidity in a study (e.g., selection bias, maturation, history, attrition). Due to lack of randomization at the time of the ACES program delivery, this study used a non-equivalent comparison group design, a frequently used type of quasi-experimental design (Bordens & Abbott, 2008; Fraenkel & Wallen, 2009; Shadish, Cook & Campbell, 2002), to address the evaluation questions. Comparison groups of students were created using multivariate statistical techniques as described below.

For Year One and Year Two analyses, this study used propensity scores matching procedures to create individual-level comparison groups via the application of multivariate analytical procedures (Dehija & Wahba, 2002). Two separate comparison groups were created for each grade level in each year. Specifically, the matched samples were drawn from Grade 11 and Grade 12 MCPS high school students with the following two conditions. First, the matched samples had similar characteristics as measured by race/ethnicity; receipt of the following services: English for Speakers of Other Languages (ESOL), Free and Reduced-price Meals System (FARMS), and special education; and, as a measure of initial abilities as measured by an academic characteristic, PSAT score, as the ACES group of students. Second, the matched samples were similar in size to the ACES student population. Matching was done using the matching package in R (Sekhon, 2006) for Year One and the matching component available through IBM SPSS Statistics software for Year Two.

For Year One analyses, this study created a second comparison group, a school-level matched group of students. A description of methods for the school-level matched comparison group can be found in Appendix B.

Data Sources

The relevant data for outcome analyses were either extracted from MCPS student level records or provided by MC (i.e., ACCUPLACER scores for students who took the test through the college).

Study Samples

Individual-level matched for Year Two. The analytical file for Year Two included 1,586 11th and 12th graders who were enrolled in MCPS during the 2014–2015 school year. The 11th grade sample included 425 ACES and 425 matched non-ACES students. The 12th grade sample included 368 ACES and 368 matched non-ACES students. Please note only ACES students who had a match and students who had information on all outcome variables were included in the final dataset. Furthermore, ACES students who were in Grade 10 based on their earned high school credits or who transferred out of MCPS were not included. Student characteristics between the two groups were similar; a table of student characteristics for both groups can be found in Appendix A, Table A1.

Individual-level matched for Year One. The analytical file for Year One included 1,558 11th and 12th graders. The Grade 11 sample included 345 ACES and 345 matched non-ACES students. The Grade 12 sample included 434 ACES and 434 matched non-ACES students. The students included in the file were enrolled in MCPS schools during the 2013–2014 school year. Please note only ACES students who had a match and students who had information on all outcome variables were included in the final dataset. Furthermore, ACES students who were in Grade 10 based on their earned high school credits or who transferred out of MCPS were not included. Student characteristics between the two groups were similar and are shown in Appendix A, Table A2.

School-level matched for Year One. Information on the study sample in Year One using the second comparison group, a school-level matched group, can be found in Appendix B.

Outcome Measures

The high-school outcome measures included the following: end-of-year grade point average (GPA), Quarter 4 marking period average (Q4 MPA), sum of unexcused absences, and scores on ACT and SAT tests.

The college-readiness outcome measures included the following: promotion to next level (i.e., promotion to 12th grade for 11th graders, graduation for 12th graders); completing SAT or ACT tests; and attainment of college-readiness milestones for SAT/ACT, for Advanced Placement or International Baccalaureate (AP/IB) exams, and for ACCUPLACER placement tests. The MCPS milestones for college and career readiness include an SAT (combined reading, mathematics, and writing) score of 1650 or higher or an ACT score of 24 or higher (MCPS, 2014). The MCPS college-readiness milestone for AP/IB exams is earning a 3 or higher on an AP exam or 4 or higher on an IB exam. The benchmark for the ACCUPLACER placement test referred to whether students met MC's minimum college-ready scores. At MC, students are exempt from remedial English courses by attaining a score of 90 or above on ACCUPLACER English and a 79 or above

on ACCUPLACER reading. The benchmark for ACCUPLACER mathematics was a score of 45 or higher; with this score, a student is eligible to take mathematics courses at MC that are college-level and credit-bearing for certain majors (e.g., humanities, arts, social sciences). Note that for other majors (e.g., business, engineering, science), MC's required score varies, but is higher than 45.

Analytical Procedures

Both statistical significance tests and effect sizes, where appropriate, were used to address the outcome questions. The former examines the likelihood that observed differences between the groups of students (i.e., ACES group vs. a comparison group) occurred by chance. Statistical significance tests, however, are influenced by sample sizes such that with a large sample, even small differences may be significant. Many researchers (e.g., American Psychological Association, 2001; Carver, 1993; Cohen, 1988; Levin, 1993; Thompson, 1995) recommend including estimates of the practical significance (e.g., the effect size) as well as the results of statistical significance tests. The effect size index is scale invariant or metric-free and can be used to interpret patterns of outcome measures' differences between groups of students and across different measures (Bloom et al., 2008; Lipsey et al., 2012). The effect size (ES) differs from statistical significance in that it measures the magnitude of the difference between two groups rather than whether the difference was due to chance. This study uses Cohen's *d* convention by which an ES with an absolute value of .2 is considered small, an ES of at least .5 is considered medium, and an ES of .8 or greater is considered large (Cohen, 1988).

Question 1

To address Question 1 on significant differences in high school outcomes between ACES and the comparison group of students, analysis of covariance (ANCOVA) and logistic regression analyses were used for both Year One and Year Two, as described below.

ANCOVA statistical procedures (Kirk, 1995) were utilized to evaluate the outcomes of the ACES program for the following six continuous outcome measures: end-of-year GPA; Q4 MPA; sum of unexcused absences; total scores for SAT reading and mathematics; total scores for SAT reading, mathematics, and writing; ACT composite score. To balance the groups of students (ACES and comparison), propensity scores (based on students' race/ethnicity, limited English proficiency status, FARMS status, and special education status) were computed (Luellen, Shadish, & Clark, 2005). The propensity scores were then divided into five categories and incorporated as categorical covariates in the statistical models (Rosenbaum & Rubin, 1983, 1984, 1985). The use of the propensity scores as covariates provided an effective avenue for controlling any preexisting differences between the groups of students resulting in less biased estimates of the ACES effects. Moreover, a measure of initial abilities of students (e.g., GPA from the previous school year, Quarter 1 MPA) also was included in each of the statistical models as a covariate where available and appropriate. Please note the pre-program measure (initial ability) is different from the post measures (outcomes) in most of the analyses presented in this study. In advanced statistical analyses, the pre-program measure can be any variable (or test) measured prior to the program intervention that is highly correlated with the post-program measure.

To control for non-parallelism or interaction, the product term between covariate and grouping variable (ACES and comparison) was included in each of the models. Six ANCOVA models (one for each outcome measure) were constructed at each grade level (11th or 12th) for a total of 12 ANCOVA models. Each of the models used one of the outcome measures as the dependent variable. The independent variable in each of those models was a dummy variable created to represent the status of the students' experience (1=ACES group and 0=comparison group). The categorical propensity score was included in each of those models to control for students' demographic and service receipt status. Effect sizes were calculated from the adjusted means provided by ANCOVA analyses to judge whether the observed differences between student groups (ACES vs. comparison) were large enough to be of practical significance to educators (Kline, 2004; Lipsey, & Freeman, 2004). The following formula was used to calculate the effect sizes (ES) from the ANCOVA findings: $ES = (M_t - M_c)/SD$. The M_t and M_c are adjusted group means for the ACES group (treatment) and the comparison group respectively, and SD is the standard deviation of the pooled outcome scores (Thompson, 2002).

Question 2

The analytical procedures to determine whether there were significant differences in college-readiness outcomes between ACES and a comparison group of students varied by year, as described below.

Year Two. A total of eight logistic regression models (two per outcome—one for Grade 11 and one for Grade 12) were used for the following four, dichotomous college-readiness outcomes: promotion to next level, took SAT or ACT, met SAT/ACT milestone, met AP/IB milestone. Two additional logistic regression models were used for Grade 12 only on the following dichotomous college-readiness outcomes: met ACCUPLACER reading/English benchmark, met ACCUPLACER mathematics benchmark. Specifically, logistic regression procedures were applied to compare the likelihood of achieving each outcome between ACES students and the comparison students, while several student characteristic variables were held constant. Odds ratios from the logistic regression models were reported to show whether or not the probability of achieving each outcome was higher for ACES students than for their peers in the comparison group. Effect sizes were calculated to estimate the magnitude of the differences in the probability of achieving each outcome. Odds ratio and Cohen's d statistics are not on the same scale; therefore, odds ratios were converted to the Cohen's effect sizes for the interpretation of findings. The following formula was used to calculate the effect size from odds ratio:

$$\text{logit } d = \frac{\ln(OR)}{\pi / \sqrt{3}}$$

Year One. For each college-readiness outcome, Fisher's Exact Test from chi-square analyses was used to detect whether differences between the ACES group and the comparison group in the percentage of students who met each outcome were statistically significant. Separate analyses were done for Grade 11 and Grade 12 students. Data on ACCUPLACER tests were not available.

Strengths and Limitations

Strengths

Evaluators used two control techniques for improving the internal validity of the study and for estimating a less biased effect of the ACES program.

Control by study design. The key component of the quasi-experimental design is the use of appropriate comparison groups when evaluating a program's outcomes. In this evaluation, multivariate analyses were used to create two comparison groups in Year One and one in Year Two.

Control by statistical techniques. Prominent researchers (Campbell & Stanley, 1963; Judd, Smith, & Kidder, 1991) argued that in order to observe the true effects of treatment in quasi-experimental designs, ANCOVA should be conducted to control for confounding variables. Therefore, ANCOVA procedures were utilized in this study to evaluate the outcomes of the ACES program for continuous outcome measures (Kirk, 1995). Logistic regression models were used to assess the ACES outcomes where the dependent measures were dichotomous (for Year Two). By using the above stated statistical models, the selection threats to the internal validity of findings were lessened by statistically controlling for preexisting differences between the groups of students.

Limitations

Although the study findings are based on a sound evaluation design as well as appropriate analyses, it should be noted that causality should not be inferred from the study due to lack of a randomized experimental design. This evaluation used a quasi-experimental design to address the evaluation questions. However, many investigators assert that randomized studies using experimental design are necessary to reliably determine program effects (Ashenfelter & Card, 1985; Barnow, 1987; Burtless & Orr, 1986;).

A major problem with employing the quasi-experimental design in educational settings is that the two groups of students have important preexisting differences that may influence their achievement/outcomes after exposure to an intervention, and this will consequently threaten the internal validity of the findings (Gay & Airasian, 2000). Isolating the effects of the ACES program on students' outcomes is not an easy task. There are many factors that also can affect students' outcomes but could not be controlled in this study due to their unavailability in a measureable format. For instance, whether non-ACES students are first generation students was unknown; ACES students provided that information on their applications for the program, but this information is not collected outside of the ACES program.

Three other caveats are relevant when interpreting the findings. First, ACES students self-report whether or not they are first generation students. Second, some students had other experiences that might have impacted their academic and college-readiness outcomes. For example, in Year Two, 4% of the Grade 11 ACES students and 6% of the Grade 12 ACES students were in the Advancement Via Individual Determination (AVID) program, which is an elective course that prepares students for college (AVID, 2015). Similarly, three (less than 1%) of the non-ACES

Grade 12 students in Year Two participated in College Tracks, a program designed to work with students to help them apply to college or technical school and seek financial aid (MCPS, 2016).

Third, because matches could not be found for each ACES student, the study did not include all ACES students. Further, the matched ACES samples differed somewhat from all ACES students. For Year Two, the matched sample had fewer Hispanic/Latino students, more White students, and fewer students who received ESOL services in a prior school year. For Year One, the matched sample had more students who were FARMS recipients and fewer students who were ESOL recipients,

Only a classical experiment with the random assignment of students can safeguard against each of the sources of internal invalidity in a study (e.g., selection bias, maturation, history, attrition) as noted by several researchers (Babbie, 1992; Judd, Smith, & Kidder, 1991).

Finally, this report excludes analyses of college outcomes, such as application, acceptance, and plans for attendance. Although these outcomes are key, this information was available only for ACES students, and not for the comparison students.

Organization of Findings

The next sections are a summary of findings followed by study conclusions. The findings are divided into two sections. The first section presents results of the study for Year Two of ACES (2014–2015). The second section presents results for Year One of ACES (2013–2014). Appendix B presents the results of Year One using school-level comparison groups, along with the methodology associated with these findings.

Findings for Year Two

This section describes findings from comparisons of students in the ACES program to students in the comparison group on several outcome measures. Analyses included only students who attended high schools in MCPS during 2014–2015 and had complete information on outcome variables, demographics, and service receipt measures. Analyses were done separately for 11th and 12th graders. Findings are displayed in the order of the evaluation questions.

Question 1. Are there significant differences in high school outcomes between ACES and a comparison group of students?

The findings for Question 1 included adjusted mean differences between ACES and the comparison group of students; these differences are calculated based on matched samples, with control and covariate variables included. The analyses can be thought of as testing whether the adjusted mean difference between the two groups (ACES vs. comparison) is significantly different from zero. There were two tests of significant differences: statistical (indicated by $p < .05$) and practical (indicated by ES with an absolute value of .20 or greater, using Cohen’s convention).

Grade 11

Table 1 presents the ANCOVA findings for Question 1 for Grade 11.

2015 GPA. The descriptive findings (Table 1, original means and standard deviations) indicated that the average GPA of the ACES group of students (2.85) was slightly higher than that of the comparison group (2.70). The ANCOVA, which controlled for the effects of 2014 GPA, demographics, and service receipt measures, did not find a statistically significant ($p > .05$) difference between the two groups of students pertaining to 2015 GPA. Moreover, the magnitude of the adjusted mean differences in 2015 GPA between the two groups of students was not of any practical significance as calculated by the effect size (ES = .02). These findings suggested that ACES Grade 11 students performed as well as comparison students as measured by 2015 GPA.

Table 1
ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores
Between ACES and a Comparison Group of Grade 11 Students in Year Two

Outcome measure	Means						ACES effect		
	ACES group (N = 425)			Comparison group (N = 425)			Adjusted mean difference	Standard error	Effect size
	N	Original mean (SD)	Adjusted mean	N	Original mean (SD)	Adjusted mean			
2015 GPA	425	2.85 (.59)	2.78	405	2.70 (.64)	2.76	0.02	0.01	.02
Q4 MPA	423	2.92 (.78)	2.88	411	2.76 (.88)	2.81	0.07	0.04	.08
Sum of unexcused absences	425	3.62 (9.64)	3.84	422	5.80 (12.24)	6.11	2.27*	0.76	.21
SAT (reading & math) score	255	960.98 (178.17)	962.53	185	979.62 (205.89)	976.30	-13.77	14.10	-.07
SAT (reading, math, & writing) score	255	1,438.04 (259.00)	1,439.72	185	1,466.97 (304.47)	1,463.36	-23.64	20.40	-.09
ACT score	100	19.32 (4.48)	19.49	46	20.89 (4.83)	20.79	-1.31	.68	-.28

Note. SD = Standard deviation.

* $p < .05$

Q4 MPA. The descriptive findings (Table 1 above) indicated, on average, the Q4 MPA among the ACES group (2.92) was higher than that of the comparison group (2.76). However, further analyses revealed that the adjusted mean difference was neither statistically ($p > .05$) nor practically (ES = .08) significant. These findings suggested that ACES Grade 11 students performed as well as comparison group of students as measured by 2015 Q4 MPA.

Sum of unexcused absences. The descriptive findings showed, on average, ACES students had a lower sum of unexcused absences (3.62) than the comparison students (5.80) (Table 1 above). The adjusted mean difference (2.27) as calculated by ANCOVA, controlling for student demographics and service receipt measures, was statistically significant ($p < .05$). The effect size associated with the adjusted mean difference also was practically significant (ES = .21), according to Cohen's convention. These results suggested that Grade 11 ACES students fared better than the comparison students as measured by the sum of unexcused absences.

SAT combined score (reading and mathematics). With respect to combined SAT reading and mathematics score, the original mean for the ACES group (960.98) was lower than that for the comparison group (979.62) (see Table 1 above). However, after controlling for students' prior performance (2014 GPA), demographics, and service receipt measures, the comparison students did not have a statistically higher mean ($p > .05$) than the ACES students (adjusted mean difference = -13.77). Also, the difference between the two groups was not practically significant (ES = -.07). These findings suggested that the comparison 11th graders performed as well as the ACES 11th graders, as measured by their combined SAT reading and mathematics scores.

SAT combined score (reading, mathematics, and writing). The descriptive findings showed that comparison students had a higher mean score on their combined SAT reading, mathematics, and writing (1,466.97) than the ACES students (1,438.04) (Table 1 above). However, ANCOVA

analyses revealed that the difference between the groups (adjusted mean difference = -23.64) was not statistically significant ($p > .05$) or meaningful in an educational setting ($ES = -.09$).

ACT composite score. The analyses (Table 1 above) revealed that there was not a statistically significant difference ($p > .05$). However, there was a small, practically significant difference ($ES = -.28$) between the two groups of Grade 11 students on adjusted ACT mean scores (-1.31), in favor of comparison students.

Grade 12

Table 2 presents the ANCOVA findings for Question 1 for Grade 12.

2015 GPA. The descriptive findings (Table 2, original means and standard deviations) indicated that the average GPA of the ACES group of students (2.73) was higher than the comparison group (2.62). ANCOVA, which controlled for 2014 GPA, demographics, and service receipt measures, did not find a statistically significant ($p > .05$) difference between the two groups for 2015 GPA. Moreover, the adjusted mean differences in 2015 GPA between the groups of students was not of any practical significance based on the effect size ($ES = .03$). These findings suggested that Grade 12 ACES students performed as well as comparison students on 2015 GPA.

Table 2
ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores
Between ACES and a Comparison Group of Grade 12 Students in Year Two

Outcome measure	Means						ACES effect		
	ACES group (N = 368)			Comparison group (N = 368)			Adjusted mean difference	Standard error	Effect size
	N	Original mean (SD)	Adjusted mean	N	Original mean (SD)	Adjusted mean			
2015 GPA	368	2.73 (.60)	2.69	368	2.62 (.67)	2.67	0.02	0.01	.03
Q4 MPA	365	2.64 (.91)	2.58	346	2.54 (.93)	2.58	0.00	0.05	.00
Sum of unexcused absences	368	6.84 (11.92)	6.62	352	8.32 (12.90)	8.43	-1.82*	0.93	-.15
SAT (reading & math) score	266	955.34 (183.02)	958.47	220	990.59 (192.37)	979.91	-21.44	13.00	-.11
SAT (reading, math, & writing) score	266	1,422.37 (267.67)	1,426.99	220	1,477.00 (285.46)	1,459.62	-32.62	18.52	-.12
ACT score	149	19.36 (4.54)	19.69	90	20.61 (5.15)	19.94	-0.25	0.49	-.05

Note. SD = Standard deviation

* $p < .05$

Q4 MPA. The descriptive findings (Table 2 above) indicated, on average, the Q4 MPA of the ACES group (2.64) was higher than that of the comparison group (2.54). However, there was no difference between the ACES and comparison group in the adjusted mean, as calculated by ANCOVA which controlled for Q1 MPA. Further analyses revealed that the adjusted mean difference was neither statistically ($p > .05$) nor practically ($ES = .00$) significant. These findings suggested that ACES Grade 12 students performed as well as the comparison group of students as measured by Q4 MPA.

Sum of unexcused absences. As shown in Table 2 (above), the descriptive findings showed, on average, ACES students had a lower sum of unexcused absences (6.84) than the comparison group of students (8.32). The adjusted mean difference (-1.82), as calculated by ANCOVA, which controlled for student demographics and service receipt measures, was statistically significant ($p < .05$), but the effect size associated with this adjusted mean difference was not practically significant ($ES = -.15$). These findings suggested that ACES students fared statistically better than the comparison students as measured by the sum of unexcused absences, but that the difference was not meaningful in an educational setting.

SAT combined score (reading and mathematics). As shown in Table 2 (above), students in the comparison group had a higher mean score on their combined SAT reading and mathematics (990.59) compared to their peers in the ACES group (955.34). After controlling for students' prior performance (2014 GPA), demographics, and service receipt measures, the comparison group of students did not have a statistically higher mean ($p > .05$) than the ACES group (adjusted mean difference = -21.44). Additionally, the performance difference between the groups of students was not practically significant ($ES = -.11$) according to Cohen's convention.

SAT combined score (reading, mathematics, and writing). Similar analyses revealed that students in the comparison group had a higher mean score on their combined SAT reading, mathematics, and writing (1,477.00) when compared to their peers in the ACES group (1,422.37) (Table 2 above). After controlling for students' 2014 GPA, demographics, and service receipt measures, the magnitude of difference between the two groups (adjusted mean difference = -32.62) was not statistically significant ($p > .05$) or meaningful in an educational setting ($ES = -.12$).

ACT composite score. The analyses (Table 2 above) revealed that there was not a statistically or practically significant ($p > .05$; $ES = -.05$) difference between the two groups of students on adjusted ACT mean scores (adjusted mean difference = -.25).

Question 2. Are there significant differences in college-readiness outcomes between ACES and a comparison groups of students?

Grade 11

Table 3 displays the Grade 11 results for college-readiness outcomes.

Promotion to next level. The results from logistic regression (odds ratio = 1.50) indicated that after controlling for students' demographics and service receipt measures, the probability (or chance) of promotion to Grade 12 was 1.5 times higher for 11th graders in the ACES group than for students in the comparison group (Table 3). Although the odds ratio was not statistically significant ($p > .05$), the effect size calculated from the odds ratio was practically significant ($ES = .23$) using Cohen's convention. These findings suggested that ACES students were promoted to Grade 12 at a higher rate when compared to their peers in the comparison group of students.

Table 3
Logistic Regression Results Comparing Promotion, SAT or ACT Participation, SAT/ACT Milestone, and AP/IB Milestone Between ACES and a Comparison Group of Grade 11 Students for Year Two

Outcome measure	Descriptive statistics by group						Logistic regression results			
	ACES (N = 425)			Comparison (N = 425)			B (SE)	Wald	Odds ratio	Effect size
	N	n	%	N	n	%				
Promotion to next level	425	386	90.8	425	365	86.7	.41 (.22)	3.34	1.50	.23
Took SAT or ACT	425	304	71.5	425	224	52.7	.66 (.15)	18.63**	1.93	.36
Met SAT/ACT milestone ^a	304	63	20.7	224	50	22.3	-.23 (.23)	1.05	0.79	-.13
Met AP/IB milestone	425	151	35.5	425	123	28.9	.12 (.16)	0.57	1.13	.07

Note. SE=standard error.

^aN = students who took the SAT or ACT test

* $p < .05$, ** $p < .01$

Took SAT or ACT. As seen in Table 3 above, the results from logistic regression (odds ratio = 1.93) indicated that after controlling for students' demographics and service receipt measures, the probability of an 11th grader in the ACES group taking the SAT or ACT was almost two times (1.9) that of an 11th grader in the comparison group. This odds ratio was statistically significant ($p < .01$) and the effect size calculated from the odds ratio was practically significant (ES = .36) using Cohen's convention. These findings indicated that ACES students in Grade 11 took the SAT or ACT at a higher rate than the comparison group of students.

Met SAT/ACT milestone. The results from logistic regression (odds ratio = 0.79) indicated that, after controlling for students' demographics and service receipt measures, the probability of an 11th grader in the ACES group meeting the SAT or ACT milestone was lower than the probability of an 11th grader in the comparison group meeting the milestone (i.e., odds ratio < 1) (Table 3 above). However, the odds ratio was not significant ($p > .05$) and the effect size, calculated from the odds ratio, was not practically significant (ES = -.13). These findings suggested that Grade 11 ACES students met the SAT or ACT milestone at a rate that was similar to the comparison group of students.

Met AP/IB milestone. The results from logistic regression (odds ratio = 1.13) indicated that after controlling for students' demographics and service receipt measures, the chance of meeting the AP/IB milestone was 1.13 times higher for the 11th graders in the comparison group compared to the ACES group (Table 3 above). The odds ratio was not statistically significant ($p > .05$); likewise, the effect size that was calculated from the odds ratio was not practically significant (ES = .07) using Cohen's convention. These findings implied that Grade 11 ACES students met the AP/IB benchmark at a similar rate as the comparison group of students.

Grade 12

Table 4 displays the Grade 12 results for college-readiness outcomes.

Promotion to next level. The results from logistic regression (odds ratio = 1.19) showed that after controlling for students' demographics and service receipt measures, the probability (or chance) of graduating among ACES students did not differ statistically ($p > .05$) or practically (ES = .09) from the probability of graduating among non-ACES students (Table 4).

Table 4
 Logistic Regression Results Comparing Promotion, SAT and ACT Participation, SAT and ACT Milestone, and AP or IB Milestone Between ACES and a Comparison Group of Grade 12 Students in Year Two

Outcome measure	Descriptive statistics by group						Logistic regression results			
	ACES (N =368)			Comparison (N =368)			B (SE)	Wald	Odds ratio	Effect size
N	n	%	N	n	%					
Promotion to next level	368	365	99.2	368	360	97.8	.17 (.84)	0.04	1.19	.09
Took SAT or ACT	368	308	83.7	368	245	66.6	.84 (.22)	14.64**	2.32	.46
Met SAT/ACT milestone ^a	308	60	19.5	245	72	29.4	-.30 (.25)	1.48	0.74	-.17
Met AP/IB milestone	368	154	41.8	368	127	34.5	.26 (.19)	2.00	1.30	.14
Met ACCUPLACER reading/English benchmark ^a	163	46	28.2	87	26	29.9	-.23 (.34)	0.43	0.80	-.12
Met ACCUPLACER math benchmark ^a	174	13	7.5	91	10	11.0	-.27 (.51)	0.27	0.77	-.15

Note. SE=standard error.

^aN = students who took the test

* $p < .05$, ** $p < .01$

Took SAT or ACT. The results from logistic regression (odds ratio = 2.32) revealed that after controlling for students' demographics and service receipt measures, the probability of a Grade 12 student in the ACES program taking the SAT or ACT was more than twice (2.3) the probability that a Grade 12 students in the comparison group took either test (Table 4 above). The odds ratio was statistically significant ($p < .01$), and the effect size calculated from the odds ratio was practically significant (ES = .46), using Cohen's convention. These findings indicated that ACES 12th graders took the SAT or ACT at a higher rate when compared to their peers.

Met SAT/ACT milestone. The results from logistic regression (odds ratio = .74) showed that the probability of test takers meeting the SAT or ACT milestone did not differ significantly ($p > .05$) or practically between the two groups of 12th graders (ES = -.17) (Table 4 above).

Met AP/IB milestone. Likewise, the results from logistic regression (odds ratio = 1.30) indicated that after controlling for students' characteristics, the chance of meeting the AP/IB milestone did not differ significantly ($p > .05$) between the groups of Grade 12 students (Table 4 above). The effect size calculated from the odds ratio was not practically significant (ES = .14).

Met ACCUPLACER reading/English benchmark. Findings (odds ratio = .80) revealed that after controlling for students' demographics and service receipt measures, the probability of meeting the ACCUPLACER reading/English benchmark was not significantly higher for the ACES group than for the comparison group among 12th grade test takers ($p > .05$) (Table 4 above). Also, the effect size calculated from the odds ratio was not practically significant (ES = -.12)

Met ACCUPLACER mathematics benchmark. The results from logistic regression (odds ratio = .77) revealed that after controlling for students' characteristics, the chance of meeting the ACCUPLACER mathematics benchmark did not differ statistically ($p > .05$) or practically (ES = -.15) between the two groups of Grade 12 test takers (Table 4 above).

Findings for Year One

The following sections address both evaluation questions for Year One of the ACES program. The analyses included only students who attended MCPS high schools in the 2013–2014 school year with complete information on outcome variables, demographics, and service receipt measures.

Question 1: Are there significant differences in high school outcomes between ACES students and a comparison group of students?

The findings for Question 1 included adjusted mean differences between ACES and the comparison group of students; these differences are calculated based on matched samples, with control and covariate variables included. These analyses can be thought of as testing whether the adjusted mean difference between the two groups of students is significantly different from zero. There were two tests of significant differences: statistical (indicated by $p < .05$) and practical (indicated by ES with an absolute value of .20 or greater, using Cohen’s convention).

Grade 11

Table 5 presents the ANCOVA findings for students in Grade 11.

2014 GPA. The descriptive findings (Table 5, original means and standard deviations) indicated that the average 2014 GPA was slightly higher for the ACES group (2.78) than for the comparison group (2.71). The ANCOVA, which controlled for the effects of 2013 GPA, demographics, and service receipt measures, did not find a statistically significant ($p > .05$) difference between the two groups of students pertaining to 2014 GPA. Moreover, the magnitude of the adjusted mean difference in 2014 GPA between the two groups of students was not of any practical significance, as indicated by the effect size (ES = .02). These findings suggested that ACES students performed as well as comparison students based on 2014 GPA for 11th graders.

Table 5
ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores
Between ACES and a Comparison Group for Grade 11 in Year One

Outcome measure	Means						ACES effect		
	ACES group (N =345)			Comparison group (N =345)			Adjusted mean difference	Standard error	Effect size
	N	Original mean (SD)	Adjusted mean	N	Original mean (SD)	Adjusted mean			
2014 GPA	345	2.78 (.57)	2.75	345	2.71 (.61)	2.74	0.01	0.01	.02
Q4 MPA	342	2.83 (.77)	2.79	329	2.77 (.83)	2.85	-0.06	0.03	-.07
Sum of unexcused absences	345	2.17 (5.44)	2.19	345	3.12 (7.81)	3.16	-0.97	0.51	-.14
SAT (reading & math) score	218	955.32 (175.68)	953.67	167	1,005.33 (206.22)	991.88	-38.27*	15.14	-.20
SAT (reading math, & writing) score	218	1,422.84 (255.64)	1,421.06	167	1,496.17 (298.83)	1,476.58	-55.52*	21.70	-.20
ACT score	80	19.22 (4.68)	19.64	54	22.19 (5.38)	22.16	-2.52*	0.64	-.49

Note. SD = standard deviation

* $p < .05$

Q4 MPA. The descriptive findings (Table 5 above) indicated, on average, the Q4 MPA among the ACES group (2.83) was higher than that of the comparison group (2.77). However, the adjusted mean difference, controlling for Q1 MPA, calculated by ANCOVA between the ACES and comparison group was .06 point higher in favor of the comparison group. Further analyses revealed that the adjusted mean difference was neither statistically ($p < .05$) nor practically ($ES = -.07$) significant. These findings suggested that, among Grade 11 students, the ACES group performed as well as the comparison group on Q4 MPA.

Sum of unexcused absences. The descriptive findings showed, on average, ACES students had a lower sum of unexcused absences (2.17) than the comparison students (3.12) (Table 5 above). The adjusted mean difference (-.97) as calculated by ANCOVA, controlling for student demographics and service receipt measures, was in favor of the ACES group because a lower sum of unexcused absences were more favorable. However, it did not reach a significance level ($p = .06$) and the effect size associated with the adjusted mean difference was not practically significant ($ES = -.14$). These findings suggested that ACES students were similar to the comparison students as measured by sum of unexcused absences for Grade 11 students.

SAT combined score (reading and mathematics). As shown in Table 5 (above), students in the comparison group had a higher mean score on their combined SAT reading and mathematics (955.32) compared to their peers in the ACES group (1005.33). After controlling for students' prior performance (2013 GPA), demographics, and service receipt measures, the comparison group had a statistically higher mean ($p < .05$) than the ACES group (adjusted mean difference = -38.27). The performance difference between the groups of students was also practically significant ($ES = -.20$) according to Cohen's convention. These findings suggested that the comparison group performed significantly (statistically and practically) better than the ACES students, as measured by 11th graders' combined SAT reading and mathematics scores.

SAT combined score (reading, mathematics, and writing). Similar analyses revealed, on average, 11th graders in the comparison group had a higher mean score on their combined SAT reading, mathematics, and writing when compared to their peers in the ACES group after controlling for students' 2013 GPA, demographics, and service receipt measures (Table 5 above). The difference between the two groups (adjusted mean = -55.52) was statistically significant ($p < .05$) and meaningful in an educational setting ($ES = -.20$) in favor of the comparison group.

ACT composite score. The analyses (Table 5 above) revealed that there was a difference between the two groups of Grade 11 students on adjusted mean ACT composite score (-2.52). The ACT composite score difference between the two groups was both statistically and practically significant ($p < .05$; $ES = -.49$) in favor of the comparison group.

Grade 12

Table 6 presents the ANCOVA findings for students in Grade 12.

2014 GPA. The descriptive findings (Table 6, original means and standard deviations) indicated that the average 2014 GPA among the ACES group of students (2.58) was higher than that of the comparison group (2.49). The ANCOVA, which controlled for 2013 GPA,

demographics, and service receipt measures, did not find statistically significant differences (adjusted mean difference = .02; $p > .05$). Also, the magnitude of the adjusted mean differences in 2014 GPA between the two groups of Grade 12 students was not practically significant ($ES = .03$). These findings suggested that ACES students were academically similar to the comparison students as measured by the 2014 GPA of 12th graders.

Table 6
ANCOVA Results Comparing GPA, Q4 MPA, Unexcused Absences, SAT Scores, and ACT Scores
Between ACES and a Comparison Group for Grade 12 in Year One

Outcome measure	Means						ACES effect		
	ACES group (N =434)			Comparison group (N =434)			Adjusted mean difference	Standard error	Effect size
	N	Original mean (SD)	Adjusted mean	N	Original mean (SD)	Adjusted mean			
2014 GPA	434	2.58 (.51)	2.54	434	2.49 (.64)	2.52	0.02	0.01	.03
Q4 MPA	432	2.57 (.83)	2.54	432	2.41 (.97)	2.49	0.05	0.05	.06
Sum of unexcused absences	434	4.17 (7.99)	4.04	432	6.00 (11.02)	5.85	-1.82*	0.50	-.19
SAT (reading & math) score	247	896.07 (155.72)	901.41	222	939.01 (172.68)	932.97	-31.56*	12.75	-.19
SAT (reading math, & writing) score	247	1,338.87 (226.92)	1,348.34	222	1,398.29 (252.24)	1,391.67	-43.32*	18.59	-.18
ACT score	159	17.54 (3.66)	18.09	92	19.92 (4.30)	19.41	-1.32*	0.49	-.32

Note. SD = Standard deviation

* $p < .05$

Q4 MPA. As shown in Table 6 (above), on average, Q4 MPA of the ACES group of students (2.57) was higher than that of the comparison group (2.41). However, the adjusted mean difference between the ACES and the comparison group was neither statistically ($p > .05$) nor practically significant ($ES = .06$). These findings suggested that ACES students were similar to the individual-level matched comparison students as measured by the Q4 MPA for Grade 12 students.

Sum of unexcused absences. On average, ACES students had a lower sum of unexcused absences (4.17) than the comparison group of students (6.00) in Grade 12 (Table 6 above). The adjusted mean difference (-1.82), calculated by ANCOVA after controlling for student demographic and service receipt measures, was statistically significant ($p < .05$) but not practically significant ($ES = -.19$). These findings suggested that ACES 12th graders had a significantly lower average sum of unexcused absences than the comparison 12th graders. The magnitude of the difference was not significant in an educational setting.

SAT combined score (reading and mathematics). Grade 12 students in the comparison group had a higher mean score on their combined SAT reading and mathematics (939.01) compared to their peers in the ACES group (896.07) (Table 6 above). After controlling for students' prior performance (2013 GPA), demographics, and service receipt measures, the comparison group of students had a significantly higher mean than the ACES students (adjusted

mean difference = -31.56; $p < .05$). However, the magnitude of the difference between the groups of students was not meaningful in an educational setting ($ES = -.19$).

SAT combined score (reading, mathematics, and writing). Similar analyses (Table 6 above) revealed, on average, students in the comparison group had a higher mean score on their combined SAT reading, mathematics, and writing when compared to their peers in the ACES group (mean difference = -59.42). After controlling for students' 2013 GPA, demographics, and service receipt measures, the difference between the two groups (adjusted mean difference = -43.32) was statistically significant ($p < .05$) in favor of the comparison group. However, the difference between the groups of students was not meaningful in an educational setting ($ES = -.18$).

ACT composite score. The analyses revealed that there were differences between the two groups of Grade 12 students on adjusted ACT mean scores (-1.32) (Table 6 above). Further analyses showed, after adjusting for students' 2013 GPA, demographics, and service receipt measures, that the mean difference in ACT composite score was both statistically and practically significant ($p < .05$; $ES = -.32$). These findings indicated that the comparison 12th graders performed significantly (statistically and practically) better than the ACES 12th graders, as measured by composite ACT score.

Question 2: Are there significant differences in college-readiness outcomes between ACES students and a comparison group of students?

Grades 11 and 12

Promotion to next level. As shown in Table 7, there was no significant difference in promotion to Grade 12 between Grade 11 ACES students and their peers (91% vs. 90%, point difference = 0.6, $p > .05$). However, a significantly higher percentage of Grade 12 ACES students graduated when compared with their peers (99% vs. 95%, point difference = 4.8, $p < .001$).

Table 7
Promotion to Next Level for ACES and a Comparison Group
by Grade Level in Year One

Promotion to next level	ACES group			Comparison group			Difference %
	N	n	%	N	n	%	
Grade 11	345	313	90.7	345	311	90.1	0.6
Grade 12	434	431	99.3	434	410	94.5	4.8***

* $p < .05$, ** $p < .01$, *** $p < .001$.

Took SAT or ACT. A significantly higher percentage of Grade 11 ACES students took the SAT or ACT than their peers (73% vs. 57%, point difference=16.5, $p < .05$) (Table 8). Likewise, a significantly higher percentage of Grade 12 ACES students took either the SAT or ACT when compared with their peers (71% vs. 58%, point difference = 12.9, $p < .05$).

Table 8
Participation in SAT or ACT for ACES and a Comparison Group
by Grade Level in Year One

Took SAT or ACT	ACES group			Comparison group			Difference
	N	n	%	N	n	%	%
Grade 11	345	253	73.3	345	196	56.8	16.5*
Grade 12	434	307	70.7	434	251	57.8	12.9*

* $p < .05$, ** $p < .01$, *** $p < .001$.

Met SAT/ACT milestone. While a significantly higher proportion of ACES students took the SAT or ACT than the comparison students, fewer ACES test takers met the SAT or ACT milestone (Table 9). The gap between the ACES 11th graders and the comparison group of students (-12.6 percentage point) was statistically significant. Likewise, the gap between the ACES 12th graders and their peers was statistically significant.

Table 9
Met SAT/ACT Milestone for ACES and a Comparison Group of Test Takers
by Grade Level in Year One

Met SAT/ACT milestone ^a	ACES group			Comparison group			Difference
	N	n	%	N	n	%	%
Grade 11	253	43	17.0	196	58	29.6	-12.6*
Grade 12	307	29	9.4	251	47	18.7	-9.3*

^aN = students who took the test.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Met AP/IB milestone. As seen in Table 10, there were no significant differences between student groups with respect to meeting the AP/IB milestone. Among 11th graders, a slightly lower percentage of ACES students than comparison students met the milestone (32% vs. 34%); the difference was not significant. Among 12th graders, a slightly higher percentage of ACES students than comparison students met the AP/IB milestone (34% vs. 31%); again, the difference was not significant.

Table 10
Met AP/IB Milestone for ACES and a Comparison Group
by Grade Level in Year One

Met AP/IB milestone ^a	ACES group			Comparison group			Difference
	N	n	%	N	n	%	%
Grade 11	345	111	32.2	345	118	34.2	-2.0
Grade 12	434	147	33.9	434	135	31.1	2.8

^aN = students who took the test.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Questions 1 and 2

The analyses addressing both evaluation questions for Year One were repeated using a school-level matched sample. The findings were somewhat consistent with the findings presented below; detailed results are in Appendix B.

Summary of Findings

Evaluation Question 1

The analyses on high school outcomes revealed no significant differences between ACES students and comparison students in both Year One and Year Two on two measures: end-of-year GPA and Q4 MPA. However, there were significant differences between ACES students and their peers on the sum of unexcused absences. In the ACES group, 11th graders had fewer unexcused absences in Year Two, and 12th graders had fewer unexcused absences in both Year One and Year Two. Lastly, there were significant differences in favor of the comparison group on mean SAT and ACT scores in Year One for both Grade 11 and Grade 12; but in Year Two, the only significant difference was on ACT scores for 11th graders, again in favor of comparison students. Further, the effect sizes from the analyses of these SAT and ACT scores decreased from Year One to Year Two for both grade levels, indicating that the differences between the ACES and comparison students narrowed.

Evaluation Question 2

Analysis of college-readiness outcomes revealed significant differences in promotion to the next level as follows: in Year One, ACES students in Grade 11 were more likely to be promoted, and in year Two, ACES students in Grade 12 were more likely to graduate. In both years, there were significant differences in participation in the SAT or ACT; more ACES than comparison students in both Grade 11 and Grade 12 took the SAT or ACT. Among students who took the SAT/ACT, the percentage of ACES students who met the MCPS milestone was significantly lower than the comparison group for both Grade 11 and Grade 12 in Year One. But in Year Two, there were no significant differences between ACES and comparison students in meeting the SAT/ACT milestone. Similarly, there were no significant differences between ACES and comparison students in meeting the AP/IB milestone in either year or in meeting the ACCUPLACER benchmarks (analyzed only for Year Two).

Study Conclusion

In conclusion, this report does not provide evidence that the ACES program improved students' grades or scores on college-readiness tests (i.e., ACT, SAT, ACCUPLACER). However, it is worth noting that these measures reflect students' decisions, achievement, coursework, and other experiences prior to their entry into the ACES program and that, during Year One, the 12th graders in ACES participated in the program for only one year. It is a positive sign that the differences between the ACES students and their peers on ACT and SAT scores decreased in the second year of program implementation.

Further, there is evidence that ACES positively impacted students' decisions while in the program. For example, it appears that ACES students were more conscientious about school attendance (i.e., had fewer unexcused absences) than their peers. Further, ACES students were more likely than their non-ACES peers to take the steps necessary for completing the ACT or SAT. This finding suggests that ACES students were more intent on attending college. Additionally, as presented in

a companion report (Wolanin & Cooper-Martin, 2016), almost all of the 12th grade ACES students in Year Two applied to, gained admission to, and reported plans to attend college. Unfortunately, we were unable to test whether ACES students were more likely than their peers to accomplish these steps, due to a lack of information about comparison students.

Finally, there was evidence for a positive impact of ACES on promotion and graduation. The strongest evidence related to high school graduation, an accomplishment that reflected students' behavior and decisions while in the ACES program, as well as in prior years, and which is essential for reaching the program's long-term goal that each ACES student earn a 4-year college degree.

Acknowledgements

The authors would like to thank Mrs. Chrisandra A. Richardson, associate superintendent, Office of Special Education and Student Services at MCPS and Mrs. Karen Callender, director of the ACES Program at MC for program support and review of the report. In addition, the authors thank Ms. Kathleen Wessman, vice president for planning and institutional effectiveness at MC, and Mr. Justin Edgar, director of the Center for Recruitment & Transfer Access, Division of Student & Academic Services at the USG; for review of the report.

The authors also thank Mr. Seong (Jeff) Jang, evaluation support specialist, Program Evaluation Unit (PEU) at MCPS, for assistance with data files and review of data tables; and Julie Wade, evaluation specialist, PEU, for her helpful suggestions and review of this report.

References

- Advancement Via Individual Determination (AVID). Retrieved November 5, 2015, from <http://www.avid.org/secondary.ashx> .
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Ashenfelter, O. & Card, D. (1985). Using the longitudinal structure of earning to estimate the effect of training programs. *Review of Economics and Statistics*, 67, 648-660.
- Babbie, E. R. (1992). *The Practice of social research* (6th Edition). Belmont, CA: Wadsworth publishing company.
- Barnett, E. A., Corrin, W., Nakanishi, A., Bork, R. H., Mitchell, C., & Sepanik, S. (2012). Preparing high school students for college: An exploratory study of college readiness partnership programs in Texas. Columbia University Academic Commons, <http://dx.doi.org/10.7916/D80R9MF9>.
- Barnow, B. (1987). The impact of CETA programs on earnings. A review of literature. *Journal of Human Resources*, 22, 157–193.
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, 1(4), 289-328.
- Bordens, K. S. & Abbott, B. B. (2008). *Research design and methods: A process approach*, 7th Edition. New York, NY: McGraw-Hill.
- Burtless, G. & Orr, L. (1986). Are classical experiments needed for manpower policy? *Journal of Human Resources*, 22, 606–639.
- Campbell, D. T. & Stanley, J. C. (1963). *Experimental and Quasi-experimental designs for research*. Chicago: Rand McNally College Publishing.
- Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education*, 61(4), 287–292.
- Cates, J. T., & Schaeffle, S. (2011). The relationship between a college preparation program and at-risk students' college readiness. *Journal of Latinos and Education*, 10 (4), 320–334.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Engberg, M. E., & Wolniak, G. C. (2010). Examining the effects of high school contexts on postsecondary enrollment. *Research in Higher Education*, (2), 132.

- Fraenkel, J. R., & Wallen, N. E. (2009). *How to design and evaluate research in education*. New York, NY: McGraw-Hill.
- Gay, L. R. & Airasian, P. W. (2000). *Educational research: Competencies for analysis and application* (6th ed.). Englewood Cliffs, N.J.: Prentice Hall.
- Hedrick, T. E., Bickman, L., & Rog, D. J. (1993). *Applied research design: A practical guide*. Newbury Park, CA: Sage Publications, Inc.
- Judd C. M., Smith, E. R. & Kidder, L. H. (1991). *Research methods in social relations*. San Francisco, CA: Holt, Rinehart and Winston, Inc.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences*. New York, NY: Brooks/Cole Publishing Company.
- Kline, R. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Levin, J. R. (1993). Statistical significance testing from three perspectives. *Journal of Experimental Education*, 61(4), 378–381.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.
- Luellen, J. K., Shadish, W. R., & Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, 29(6), 530–558.
- Montgomery County Public Schools. (2015). *What does it mean to receive special education services?* Retrieved November 24, 2015, from <http://www.montgomeryschoolsmd.org/departments/special-education/common-questions/overview.aspx#overview3>.
- Montgomery County Public Schools. (2016). *College Tracks Release Form*. Retrieved January 11, 2016, from [http://www.montgomeryschoolsmd.org/uploadedFiles/schools/wheatonhs/departments/WHS%20RELEASE%20FORM%20Eng%20and%20Span%202.9.15\(1\).pdf](http://www.montgomeryschoolsmd.org/uploadedFiles/schools/wheatonhs/departments/WHS%20RELEASE%20FORM%20Eng%20and%20Span%202.9.15(1).pdf).
- Montgomery County Public Schools, Montgomery College, and the Universities at Shady Grove. (2013). *Memo of Understanding*. Unpublished Paper. Rockville, MD: Montgomery County Public Schools, Montgomery College, and the Universities at Shady Grove.
- Rossi. P. & Freeman, H. E. (1993). *Evaluation: A systematic approach*. Beverly Hills, CA: Sage. Sage Publications.

- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–45.
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, *79*, 561–524.
- Rosenbaum, P.R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling that incorporates the propensity score. *The American Statistician*, *39*, 33–38.
- Rossi, P. H., Lipsey, Mark W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: SAGE Publications, Inc.
- Sekhon, J. S. (2006). Matching: Algorithms and software for multivariate and propensity score matching with balance optimization via genetic search. Downloaded from: <http://sekhon.berkeley.edu/matching/>.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.
- Thompson, B. (1995). Editorial policies regarding statistical significance testing: Three suggested reforms. Paper presented at the annual meeting of the Mid-South Educational Research Association, Biloxi, MS.
- Thompson, B. (2002). “Statistical,” “practical,” and “clinical”: How many kinds of significance do counselors need to consider? *Journal of Counseling & Development*, *80*(1), 64–71.
- Wolanin, N. (2015). *Formative evaluation of Achieving Collegiate Excellence and Success Program at Montgomery County Public Schools Year Two: Student and parent experiences*. Rockville, MD: Montgomery County Public Schools.
- Wolanin, N. & Cooper-Martin, S. (2016). *Outcome evaluation of the Achieving Collegiate Excellence and Success Program at Montgomery County Public Schools: Year Two*. Rockville, MD: Montgomery County Public Schools.
- Wolanin, N. & Modarresi, S. (2015). *Evaluation of Achieving Collegiate Excellence and Success Program: Student outcomes Year One, Grades 11 and 12*. Rockville, MD: Montgomery County Public Schools.

Appendix A: Characteristics of ACES and Comparison Students

Table A1
 Characteristics of All ACES Students, Matched ACES Students, and Individual-level Matched Comparison Students for Year Two

Characteristics	Total ACES N = 1,019		Matched ACES ^a N = 793		Comparison students N = 793	
	n	%	n	%	n	%
Grade						
Grade 12	455	44.7	368	46.4	368	46.4
Grade 11	564	55.3	425	53.6	425	53.6
Gender						
Female	628	61.6	482	60.8	463	58.4
Male	391	38.4	311	39.2	330	41.6
Race/ethnicity						
Asian	94	9.2	72	9.1	97	12.2
Black or African American	448	44.0	348	43.9	316	39.8
Hispanic/Latino	385	37.8	269	33.9	269	33.9
White	59	5.8	93	11.7	93	11.7
Two or More Races	31	3.0	17	2.1	17	2.1
Services received						
Current FARMS	573	56.2	420	53.0	417	52.6
Current special education	80	7.9	62	7.8	37	4.7
Current ESOL	62	6.1	20	2.5	27	3.4
Prior ESOL	411	40.3	318	40.1	297	37.5
	Mean ^b	SD ^b	Mean	SD	Mean	SD
PSAT	119.5	25.3	122	24.9	122	26.3

^aExcludes ACES students who were in Grade 10 based on their earned high school credit, transferred out of MCPS, had missing variables (e.g., PSAT), or did not have a non-ACES match.

^bCalculated only for students with scores; n = 922.

Note. SD = standard deviation.

Table A2
 Characteristics of All ACES Students, Matched ACES Students, and Individual-level Matched Comparison Students for Year One

Characteristics	Total ACES N = 968		Matched ACES ^a N = 779		Comparison students N = 779	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Grade						
Grade 12	562	58.1	434	55.7	434	55.7
Grade 11	406	41.9	345	44.3	345	44.3
Gender						
Female	570	58.9	469	60.2	468	60.1
Male	398	41.1	310	39.8	311	39.9
Race/Ethnicity						
Asian	79	8.2	66	8.5	65	8.3
Black or African American	414	42.8	329	42.2	339	43.5
Hispanic/Latino	379	39.2	302	38.8	307	39.4
White	73	7.5	65	8.3	54	6.9
Two or More Races	21	2.2	15	1.9	13	1.7
Services received						
Current FARMS	492	50.8	446	57.3	429	55.1
Current special education	95	9.8	78	10.0	70	9.0
Current ESOL	72	7.4	33	4.2	21	2.7
Prior ESOL	206	21.3	168	21.6	145	18.6
	Mean ^b	SD ^b	Mean	SD	Mean	SD
PSAT	114	21.96	114	21.96	115	22.52

^aExcludes ACES students who were in Grade 10 based on their earned high school credit, transferred out of MCPS, had missing variables (e.g., PSAT), or did not have a non-ACES match.

^bCalculated only for students with scores; n = 779.

Note. SD = standard deviation.

Appendix B: Methodology and Findings for School-level Matching Comparison Group in Year One

Methodology

Design

As a first step in creating a school-level matching comparison group, the MCPS *Schools at a Glance* data for three school years (2011–2012, 2012–2013, and 2013–2014) were used to identify similar high schools. The *Schools at a Glance* provides school-level information pertaining to school and student characteristics such as: enrollment, attendance, and outcome measures at school level. The school-level matching was done via the following two multivariate analytical techniques: factor analyses and cluster statistical technique, as described below.

First, the factor analytical procedures were applied to a number of indicators (measures) from *Schools at a Glance* files for the last three years. The purpose was to produce a small number of factors. The school level indicators included in the analyses were % FARMS; % Black or African American; % Hispanic/Latino; % White; % Asian; % ESOL recipients; % special education recipients; total enrollment; % average mobility; 4-year cohort drop-out rate; 4-year cohort graduation rate; Algebra completion with a C or better by 9th grade. The factor extraction method used was principal component with an orthogonal (Varimax) rotation. Kaiser's criterion was used to determine the number of factors to be extracted from the indicators (Kaiser, 1960). After factor extraction, the analyses used regression techniques to calculate the scores on the indicators for each of the rotated factors. The factor scores were saved in the data set for the purpose of cluster analysis.

Second, a statistical technique known as cluster analysis was applied to the orthogonal factor scores to create several groups or clusters of highly similar schools. The first step in the analyses was to test different numbers of clusters to yield sufficient and similar numbers of schools in each of the clusters. The second step was to come up with the best clustering scheme (optimum number of clusters for the purpose of this study). The third step was to select a few non-ACES schools from each of the clusters. For example, if two ACES schools ended up in cluster 1, one or two non-ACES schools (depending on the number of clusters) would be selected at random from cluster 1. Using the above steps, similar schools from each of the clusters were chosen and then the 11th and 12th grade students from those schools were extracted from the MCPS data system to be used as a pool for selecting comparison students. Each cluster of schools had relatively similar average values for each of the indicators/measures included in the analyses. Finally, the following three non-ACES comparison high schools were selected for analyses: Magruder, Paint Branch, and Springbrook. The total number of 11th and 12th grade students in those three schools was 2,550 (Magruder, n=827; Paint Branch, n=894; and Springbrook, n=829). One thousand students were randomly selected from those 2,550 students (Magruder, n=326; Paint Branch, n=350; and Springbrook, n=324) so that both groups of students had similar numbers of students (ACES, n=940; school-level matched comparison, n=1,000). The demographic characteristics and service receipt status of the two groups are presented in the next section.

Study Sample

The analytical file included a total of 1,940 11th and 12th graders across the 10 ACES schools and 3 matched comparison schools. The 11th grade sample included 378 ACES and 484 matched non-ACES students. The Grade 12 sample included 562 ACES and 516 non-ACES students. Please note only students who had information on all variables were included in the final dataset. Table B1 shows characteristics of students for both matched groups and for all ACES students.

Table B1
Characteristics of All ACES Students, Matched ACES Students, and School-level Matched Comparison Students for Year One

Characteristics	Total ACES N = 968		Matched ACES ^a N = 940		School-level matched group N = 1,000	
	n	%	n	%	n	%
Grade						
Grade 12	562	58.1	562	59.8	516	51.6
Grade 11	406	41.9	378	40.2	484	48.4
Gender						
Female	570	58.9	555	59.0	484	48.4
Male	398	41.1	385	41.0	516	51.6
Race/Ethnicity						
Asian	79	8.2	76	8.1	163	16.3
Black or African American	414	42.8	402	42.8	383	38.3
Hispanic/Latino	379	39.2	370	39.4	251	25.1
White	73	7.5	71	7.6	168	16.8
Two or More Races	21	2.2	19	2.0	33	3.3
Services received						
Current FARMS	492	50.8	542	57.7	351	35.1
Current special education	95	9.8	92	9.8	94	9.4
Current ESOL	72	7.4	68	7.2	24	2.4
Prior ESOL	562	58.1	201	21.4	144	14.4
	Mean	SD	Mean	SD	Mean	SD
PSAT ^b	114	21.96	118	22.98	120	24.17

^aExcludes ACES students who were in Grade 10 based on their earned high school credit, transferred out of MCPS, or had missing variables.

^bCalculated only for students with scores; n = 779.

Note. SD = standard deviation.

Analytical Procedures

Both statistical significance tests and effect sizes, where appropriate, were used to address the outcome questions. The former examines the likelihood that observed differences between the groups of students (i.e., ACES group vs. a comparison group) occurred by chance. Statistical significance tests however, are influenced by sample sizes such that with a large sample, even small differences may be significant. Many researchers (e.g., American Psychological Association, 2001; Carver, 1993; Cohen, 1988; Levin, 1993; Thompson, 1995) recommend including estimates of the practical significance (e.g., the effect size) as well as the results of statistical significance tests. The effect size index is scale invariant or metric-free and can be used to interpret patterns of outcome measures' differences between groups of students and across different measures (Bloom et al., 2008; Lipsey et al, 2012). The effect size (ES) differs from

statistical significance in that it measures the magnitude of the difference between two groups rather than whether the difference was due to chance. This study uses Cohen's *d* convention by which an ES of 0.2 is considered small, an ES of at least 0.5 is considered medium, and an ES of 0.8 or greater is considered large (Cohen, 1988).

Question 1. To address Question 1 on significant differences in high school outcomes between ACES and the comparison group of students (i.e., school-level matched), analysis of covariance (ANCOVA) and logistic regression analyses were used, as described below.

ANCOVA statistical procedures (Kirk, 1995) were utilized to evaluate the outcomes of the ACES program for the following six continuous outcome measures: end-of-year GPA; Q4 MPA; sum of unexcused absences; total scores for SAT reading and mathematics; total scores for SAT reading, mathematics, and writing; ACT composite score. To balance the groups of students (ACES and comparison), propensity scores (based on students' race/ethnicity, limited English proficiency status, FARMS status, and special education status) were computed (Luellen, Shadish, & Clark, 2005). The propensity scores were then divided into five categories and incorporated as categorical covariates in the statistical models (Rosenbaum & Rubin, 1983, 1984, 1985). The use of the propensity scores as covariates provided an effective avenue for controlling any preexisting differences between the groups of students resulting in less biased estimates of the ACES effects. Moreover, a measure of initial abilities of students (e.g., GPA from the previous school year, Quarter 1 MPA) also was included in each of the statistical models as a covariate where available and appropriate. Please note the pre-program measure (initial ability) is different from the post measures (outcomes) in most of the analyses presented in this study. In advanced statistical analyses, the pre-program measure can be any variable (or test) measured prior to the program intervention that is highly correlated with the post-program measure.

To control for non-parallelism or interaction, the product term between covariate and grouping variable (ACES and comparison) was included in each of the models. Six ANCOVA models (one for each outcome measure) were constructed at each grade level (11th or 12th) for a total of 12 ANCOVA models. Each of the models used one of the outcome measures as the dependent variable. The independent variable in each of those models was a dummy variable created to represent the status of the students' experience (1=ACES group and 0=comparison group). The categorical propensity score was included in each of those models to control for students' demographic and service receipt status. Effect sizes were calculated from the adjusted means provided by ANCOVA analyses to judge whether the observed differences between student groups (ACES vs. comparison) were large enough to be of practical significance to educators (Kline, 2004; Lipsey, & Freeman, 2004). The following formula was used to calculate the effect sizes (ES) from the ANCOVA findings: $ES = (M_t - M_c) / SD$. The M_t and M_c are adjusted group means for the ACES group (treatment) and the comparison group respectively, and SD is the standard deviation of the pooled outcome scores (Thompson, 2002).

Question 2. For each college-readiness outcome, Fisher's Exact Test from chi-square analyses was used to detect whether differences between the ACES group and the comparison group in the percentage of students who met each outcome were statistically significant. Separate analyses were done for Grades 11 and 12.

Findings for Students in the ACES Program vs. Students in the School-level Matched Group

The following findings reflect comparisons of students in the ACES program to students in the school-level matched group. The analyses included only students who attended high schools in MCPS during 2013–2014 with complete information on outcome variables, demographics, and service receipt measures. Results are displayed in the order of the evaluation questions.

Question 1. Are there significant differences in high school outcomes between ACES and a comparison group of students?

Findings for Question 1 are presented by grade level, because analyses were done separately for 11th and 12th graders.

Grade 11

Table B2 presents the 11th grade findings from ANCOVA on high school outcomes. Analyses of findings for each of the outcome measures are presented below.

2014 GPA. The descriptive findings (Table B2) indicated that the average 2014 GPA of 11th graders was higher for ACES students (2.76) than for the school-level matched group (2.68). The ANCOVA, which controlled for the effects of 2013 GPA, demographics, and service receipt measures, was statistically significant (adjusted mean difference = .03; $p < .05$). However, the magnitude of the adjusted mean difference in 2014 GPA between the two groups of students was not of any practical significance ($ES = .04$). These findings suggested that ACES students earned a higher 2014 GPA than the comparison students, but the difference in GPA was too small to be meaningful in an educational setting.

Table B2
Comparisons of GPA, Quarter 4 MPA, Unexcused Absences, SAT Scores, and ACT Scores
Between ACES and School-level Matched Group for Grade 11 in Year One

Outcome measure	Means						ACES effect		
	ACES group (N = 378)			School-level matched group (N = 484)			Adjusted mean difference	Standard error	Effect size
	N	Original mean (SD)	Adjusted mean	N	Original mean (SD)	Adjusted mean			
2014 GPA	372	2.76 (.58)	2.73	465	2.68 (.72)	2.70	.03*	.01	.04
Q4 MPA	373	2.82 (.78)	2.79	438	2.80 (.85)	2.82	-.03	.04	-.03
Sum of unexcused absences	378	2.32 (5.42)	2.16	484	3.25 (7.24)	3.76	-1.60*	.456	-.24
SAT (reading & math) score	225	955.00 (174.04)	964.70	243	1,042.71 (190.84)	1,004.58	-39.89*	13.97	-.21
SAT (reading, math, & writing) score	225	1,423.20 (253.18)	1,437.94	243	1,544.66 (278.62)	1,489.58	-51.65*	19.89	-.19
ACT score	86	19.13 (4.56)	19.39	51	21.29 (4.92)	20.61	-1.23	.67	-.25

Note. SD = standard deviation.

* $p < .05$.

Q4 MPA. The descriptive findings (Table B2 above) indicated that, on average, the Q4 MPA for students in Grade 11 was slightly higher among the ACES group of students (2.82) than among the school-level matched comparison group (2.80). However, the adjusted mean difference between the ACES and comparison group was .03 point higher in favor of the comparison group. The adjusted mean difference was neither statistically ($p > .05$) nor practically significant ($ES = -.03$). These findings suggested that ACES students performed as well as school-level matched students as measured by their Q4 MPA.

Sum of unexcused absences. The descriptive findings for Grade 11, shown in Table B2 (above), revealed, on average, ACES students had a lower sum of unexcused absences (2.32) than the school-level matched group of students (3.25). The adjusted mean difference (-1.59) was both statistically ($p < .05$) and practically ($ES = -.24$) significant. In this case, the negative difference was in favor of the ACES group because a lower sum of unexcused absences was more favorable. These findings suggested that ACES students' unexcused absences were significantly (both statistically and practically) lower than their peers in the school-level matched group.

SAT combined score (reading and mathematics). As shown in Table B2 (above), students in the school-level matched group had a higher mean on their combined SAT reading and mathematics score (1042.71) compared to their peers in the ACES group (955.00). After controlling for students' prior performance (2013 GPA), demographics, and service receipt measures, the school-level matched group of students had a significantly higher mean score than the ACES students (adjusted mean difference = -39.89; $p < .05$). The difference between the two groups of students also was significant in an educational setting ($ES = -.21$). These findings suggested that the school-level matched group performed significantly better than the ACES group as measured by their combined SAT reading and mathematics score.

SAT combined score (reading, mathematics, and writing). Similar analyses revealed similar results for the combined SAT reading, mathematics, and writing score: students in the school-level matched group had a higher mean score (1,544.66) than their peers in the ACES group (1,423.20). Further analyses showed that the adjusted mean difference between the two groups (-51.65) was statistically significant ($p < .05$) in favor of the school-level matched group (Table B2 above). The difference between the two groups of students was not practically significant ($ES = -.19$).

ACT composite score. The analyses indicated that there were differences between the two groups of students on both unadjusted (difference = -2.16) and adjusted ACT mean scores (-1.23) (Table B2 above). The adjusted mean difference was not statistically significant ($p > .05$); nevertheless, the magnitude of the adjusted mean difference, as measured by the effect size ($ES = -.25$), was practically significant in favor of the school-level matched group.

Grade 12

Table B3 presents the Grade 12 results. The findings associated with each of the measures are described below.

2014 GPA. The descriptive findings (Table B3) indicated that the average 2014 GPA among the ACES group of students (2.57) was lower than that of the school-level matched group

(2.65). The ANCOVA analyses which controlled for the effects of possible preexisting differences, including their 2013 GPA, did not find a significant difference between the two groups of students on their 2014 GPA (adjusted mean difference = .01; $p > .05$). Moreover, the magnitude of the adjusted mean differences in 2014 GPA between the two groups of students was not practically significant (ES = .02). These findings suggested that ACES students' performance on 2014 GPA was similar to the school-level matched group.

Table B3
Comparisons of GPA, Quarter 4 MPA, Unexcused Absences, SAT Scores, and ACT Scores
Between ACES and Matched Groups of Students for Grade 12 in Year One

Outcome measure	Means						ACES effect		
	ACES group (N = 562)			School-level matched group (N = 516)			Adjusted mean difference	SE	Effect size
	N	Original mean (SD)	Adjusted mean	N	Original mean (SD)	Adjusted mean			
2014 GPA	555	2.57 (.54)	2.62	502	2.65 (.71)	2.60	.01	.01	.02
Q4 MPA	558	2.56 (.84)	2.58	480	2.65 (.92)	2.61	-.03	.04	-.04
Sum of unexcused absences	562	4.37 (7.85)	3.98	516	3.71 (7.89)	4.02	-.04	.49	-.01
SAT (reading & math) score	296	881.76 (158.12)	904.27	328	1,027.76 (201.28)	983.09	-78.82*	11.89	-.40
SAT (reading, math, & writing) score	296	1,317.65 (230.62)	1,351.45	328	1,526.24 (294.89)	1,459.78	-108.33*	17.19	-.38
ACT score	189	17.41 (3.59)	17.87	88	22.26 (5.28)	20.58	-2.71*	.48	-.57

Note. SD = standard deviation. SE = standard error.

* $p < .05$

Q4 MPA. As shown in Table B3 above, on average, Q4 MPA of the ACES group of students (2.56) was lower than that of the school-level matched group (2.65). However, the adjusted mean difference between the ACES and the school-level matched group (-.03) was neither statistically ($p > .05$) nor practically significant (ES = -.04). These findings suggested that ACES 12th graders were similar to the school-level matched students as measured by Q4 MPA.

Sum of unexcused absences. As shown in Table B3 above, on average, ACES students had a higher sum of unexcused absences (4.37) than the school-level matched group of students (3.71). The adjusted mean difference (-.04) as calculated by ANCOVA after controlling for student demographic and service receipt measures was neither statistically ($p > .05$) nor practically significant (ES = -.01). These findings suggested that ACES students in Grade 12 were similar to the school-level matched students as measured by sum of unexcused absences.

SAT combined score (reading and mathematics). Students in the school-level matched group had a higher mean score on their combined SAT reading and mathematics (1,027.76) compared to their peers in the ACES group (881.76) (Table B3 above). After controlling for students' prior performance (2013 GPA), demographics, and service receipt measures, the school-level matched students had a significantly higher mean than the ACES students (adjusted mean

difference = -78.82; $p < .05$). The magnitude of difference between the two groups of 12th graders also was meaningful in an educational setting (ES = -.40).

SAT combined score (reading, mathematics, and writing). Similar descriptive analyses revealed, on average, Grade 12 students in the school-level matched group had a higher mean score on their combined SAT reading, mathematics, and writing when compared to their peers in the ACES group (mean difference = -208.59) (Table B3 above). After controlling for students’ 2013 GPA, demographics, and service receipt measures, the difference between the two groups of students (adjusted mean difference = -108.33) was both statistically ($p < .05$) and practically significant (ES = -.38) in favor of the school-level matched group.

ACT composite score. The analyses (Table B3 above) revealed that there were differences between the two groups of Grade 12 students on both unadjusted (difference = -4.85) and adjusted ACT mean scores (difference = -2.71). Moreover, the ANCOVA findings and effect size measure revealed that the adjusted mean difference was both statistically and practically significant ($p < .05$; ES = -.57) in favor of the school-level matched group, suggesting that the school-level matched group performed significantly (both statistically and practically) better than the ACES students as measured by their composite ACT score.

Question 2. Are there significant differences in college-readiness outcomes between ACES and a comparison group of students?

Promotion to next level. As seen in Table B4, a significantly higher percentage of Grade 11 ACES students were promoted to Grade 12, compared to 11th graders in the school-level matched group (91% vs. 85%; point difference = 5.5; $p < .05$). Likewise, a significantly higher percentage of Grade 12 ACES students graduated, compared to 12th graders in the school-level matched group (99% vs. 95%; point difference = 4.6; $p < .001$).

Table B4
Promotion to Next Level for ACES and a Comparison Group
by Grade Level in Year One

Promotion to next level	ACES group			School-level matched group			Difference
	N	n	%	N	n	%	%
Grade 11	373	339	90.9	458	391	85.4	5.5*
Grade 12	561	556	99.1	511	483	94.5	4.6***

* $p < .05$, ** $p < .01$, *** $p < .001$.

Took SAT or ACT. As shown in Table B5, a significantly higher percentage of Grade 11 students in the ACES program took either the SAT or ACT than students in the school-level matched group (71% vs. 55%; point difference = 16.4; $p < .05$). For 12th graders, a slightly higher percentage of the school-level matched group took the SAT or ACT (67% vs. 68%); the difference was not significant (Table B5).

Table B5
Participation in SAT or ACT for ACES and a Comparison Group
by Grade Level in Year One

Took SAT or ACT	ACES group			School-level matched group			Difference
	N	n	%	N	n	%	%
Grade 11	378	268	70.9	484	264	54.5	16.4*
Grade 12	562	379	67.4	516	351	68.0	-0.6

* $p < .05$, ** $p < .01$, *** $p < .001$.

Met SAT/ACT milestone. As shown in Table B6, among 11th graders who took the SAT or ACT, a significantly higher percentage of the school-level matched group met the milestone compared to the ACES students (37% vs. 17%; $p < .05$). Likewise, for students in Grade 12 who took one of the tests, a significantly higher percentage of the school-level matched group met the milestone compared to the ACES students (36% vs. 8%; $p < .05$).

Table B6
Met SAT/ACT Milestone for ACES and a Comparison Group of Test Takers
by Grade Level in Year One

Met SAT/ACT milestone	ACES group			School-level matched group			Difference
	N	n	%	N	n	%	%
Grade 11	268	45	16.8	264	97	36.7	-19.9*
Grade 12	379	31	8.2	351	126	35.9	-27.7*

Note: N=students who took the test.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Met AP/IB milestone. As shown in Table B7, a significantly higher percentage of the Grade 12 school-level matched group met the MCPS AP/IB milestone compared to the ACES group (29% vs. 21%; $p < .01$). Students in the school-level matched group for Grade 11 also had a higher percentage who met the AP/IB milestone (31% vs. 26%); however, the difference was not statistically significant ($p > .05$).

Table B7
Met AP/IB Milestone for ACES and a Comparison Group
by Grade Level in Year One

Met AP/IB milestone	ACES group			School-level matched group			Difference
	N	n	%	N	n	%	%
Grade 11	378	98	25.9	484	152	31.4	-5.5
Grade 12	562	116	20.6	516	148	28.7	-8.1**

Note: N=students who took at least one AP or IB exam.

* $p < .05$, ** $p < .01$, *** $p < .001$.