WISCONSIN HOPE LAB

Evaluation of College Possible Milwaukee

June 2018

Prepared by Jed Richardson, Daniel Marlin, Jennifer Vadas, Emily Colo, Sara Goldrick-Rab

Table of Contents

Executive Summary

College Possible helps low-income students earn college degrees. They begin working with students their junior year of high school and stay with them all the way through college completion.

In the high school program, students meet with a near-peer coach for two hours, twice a week, to learn about the college search and application processes, financial aid, and ACT or SAT preparation. During their senior year, College Possible coaches help students complete applications for college, financial aid, and scholarships, and prepare for the transition to college. College Possible support with financial aid, developing study skills, and encouraging connections to local campus opportunities remains available to students throughout college. This evaluation examines the impact of College Possible for two cohorts of students attending Milwaukee Public Schools by comparing program participants to other students who have similar measurable characteristics.

College Possible aims to improve students' odds of enrolling in college and affect where they enroll, while also increasing year-to-year persistence in higher education. The program's impacts on those outcomes are examined in this report, along with additional information on students' college admissions activities and academic preparation in high school.

We estimate that students served by College Possible in Milwaukee were 18 percentage points more likely than their similar peers to enroll in college the year following high school graduation. They were also more likely to enroll in four-year and selective institutions. We find limited evidence that College Possible boosts the odds of persistence in college from the first to second year, although these estimates have wide confidence intervals because persistence outcomes currently exist for only one of the two sample cohorts.

Since students must go through an application process to join College Possible, it is possible that some of these effects are due to who participates in the program, rather than due to the program's activities. However, we find evidence that College Possible alters what students do in high school in ways that would seem to promote greater college attainment. For example, compared to their peers, College Possible students submitted more admissions and scholarship applications and had higher rates of FAFSA completion. The program also appears to have improved ACT Writing and English scores and increased enrollment in Science, Math, and AP or IB courses during twelfth grade. There is also limited evidence that the program increased the odds of high school graduation, though it did not appear to improve students' GPA, attendance, or behavior.

As expected, students who received more support from College Possible benefited more than students who received less. Students with high levels of program participation (i.e. who attended at least 50% of available program sessions) had higher rates of immediate college enrollment and enrollment in four-year colleges. They met program targets for numbers of college admissions applications and scholarship applications at much higher rates than those with lower levels of participation.

College Possible Milwaukee

College Possible aims to increase college access and success among low-income students by supporting students from low-income families in the "academic middle" as they apply to and prepare for college in their junior and senior years of high school. The program operates in six urban centers: Minneapolis/Saint Paul, MN; Milwaukee, WI; Chicago, IL; Portland, OR; Omaha, NE; and Philadelphia, PA. It also offers ongoing support to students as they pursue degrees at colleges and universities nationwide.

In Milwaukee, where the education landscape includes public, charter and voucher schools, College Possible serves low-income juniors and seniors in all three types. Rising juniors apply to the program and are admitted based on a combination of need and merit. Applicants must have family income at or below 200% of the federal poverty line or below Milwaukee county's median salary and a GPA of 2.0 or higher. Applications are also judged based on the quality of essays, teacher recommendations, and in-person interviews.

Admitted students attend program sessions with a near-peer coach for two-hours, twice a week, for the duration of their junior and senior years of high school. These near-peer coaches are drawn from the ranks of the AmeriCorps program. During the junior year students are introduced to college search and application processes; coached in the basics of college financing and financial aid; and helped to prepare for the ACT or SAT College entrance exams. They are also coached in how to use the summer months and their senior year of high school to prepare well for college. During the senior year, students work with coaches to complete college admissions applications, including drafting quality college essays and preparing for interviews. Additionally, the senior year curriculum has a strong emphasis on financial literacy and planning. Coaches assist students in completing the Free Application for Federal Student Aid (FAFSA) and scholarship applications, as well as help students anticipate the transition to college and the potential cultural and relational challenges it might present. During the summer following graduation, College Possible supports students as they complete the final steps necessary to enroll in college. Finally, programming support, such as help with financial aid, study skills, and connections to local campus resources, is provided to students throughout college as they progress toward a postsecondary degree.

Since College Possible's expansion to Milwaukee in 2008, the program has been extending its reach and increasing the number of students served in the area. In the 2016-17 school year College Possible Milwaukee served slightly fewer than 800 students in 14 partner high schools and over 1,200 college students.

Evaluation Framework

This evaluation, commissioned by College Possible Milwaukee, seeks to understand whether and to what extent College Possible impacts college enrollment, college choice, "match", and persistence among students in Milwaukee Public Schools (MPS). Impacts on students in Milwaukee's private schools are beyond the scope of this study.

We also examine the impact of College Possible on students' college admissions and preparation activities. College admissions activities include: college admissions applications, scholarship applications, and FAFSA completion. Available measures of college preparation include ACT scores and the number of college preparatory classes (math, science, and Advanced Placement (AP) or International Baccalaureate (IB) courses taken during students' senior year. Finally, the evaluation explores whether College Possible is also impacting high school engagement and performance. While such outcomes are not the express purpose of College Possible, the program could have intermediate impacts on high school engagement, leading to higher achievement and greater college readiness. High school performance outcomes include graduation as well as GPA, attendance, and disciplinary actions during grades 11 and 12.

Since participation in College Possible is voluntary and requires an application, participants differ from non-participants. Subsequent differences in their educational outcomes could be due to pre-existing differences, or due to the program itself. To distinguish between the two, we employ a quasi-experimental method called propensity score (PS) analysis. To do this, we identified a group of non-participants who are similar to College Possible students on observable characteristics, both fixed characteristics (e.g. gender) and time-variant characteristics (e.g. GPA) measured prior to program participation. These observationally similar non-participants then serve as a comparison group when estimating impacts. For more information on methodology, please see the Technical Appendix.

All MPS students who applied to and were accepted into College Possible in the 2013-14 and 2014-15 school years are included in this evaluation. In those years, College Possible operated in the following MPS high schools:

1. Hamilton High School
2. Milwaukee High School of the Arts
3. Milwaukee School of Languages
4. Morse Marshall High School
5. Pulaski High School
6. Riverside University High School
7. South Division High School
8. Hmong American Peace Academy

In addition to students served by College Possible, the evaluation also includes all students who entered the 11th grade in 2013-14 and 2014- 2015 at any of the eight MPS high schools where College Possible operated. Subsets of those students were employed for comparison purposes to estimate program impacts.

Data

Examining the educational trajectories of College Possible students and identifying similar peers for point of comparison requires good and reliable data. Program records indicated who applied to and participated in College Possible, and for how long during 11th and/or 12th grades. Students' demographic characteristics, academic performance, ACT scores, behavior, course enrollment, and survey responses from the MPS administered Senior Exit Survey regarding college preparation activities, as well college enrollment information, came from MPS and the National Student Clearinghouse (NSC), with some supplementary information from the College Possible application. College-level data on selectivity are from the Integrated Postsecondary Education Data System (IPEDS).

- *Demographic characteristics* include gender, race, English language learner status, and free or reduced price lunch participation. College Possible also provided a student-reported measure of parental education, indicating the highest level of education obtained by either parent. This measure supplemented data on the same item from the Senior Exit Survey (see below).
- *Academic performance data* include year-end cumulative GPA and credits earned for four academic years, all dates of enrollment and withdrawal, data on all courses taken during high school, days enrolled and total absence days for each year, and high school graduation.
- *ACT Scores* are available for all students and both composite and subscores are used.
- *Behavior* includes all reported disciplinary actions taken against students by year, including the full range of severity; from actions as minor as verbal correction to serious actions such as suspension or expulsion.
- *Senior Exit Survey data* include information on the number of family and friends who attended college, and for those who planned to attend college information on the number of scholarship applications submitted, number of college admissions applications submitted, and FAFSA completion.
- *College enrollment data* include whether students enrolled in college and the type of institution attended.[1]
- *College-level data include college acceptance rates.* For open-access institutions, IPEDS does not contain acceptance rates. For these, we assume rates of 100%.

[1] Although NSC includes over 3,600 institutions that enroll 98% of all students, U.S. military academies, most tribal colleges, and many very small institutions do not participate. (In 2011, NSC data included 96.5% of Wisconsin colleges, and coverage has likely increased since. See Dynarski, S., Hemelt, S. W., & Hyman, J. M. (2015). The missing manual: Using National Student Clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis, 37*(Issue 1 Supplement), 3S-5S.) In addition, colleges often do not report undocumented students. Due to these limitations, students at non-NSC colleges and undocumented students are coded as not having attended college. National Student Clearinghouse Research Center. (2014). Using NSC StudentTracker for high schools reports: considerations for measuring the college enrollment rates of high school graduates.Retrieved from https://nscresearchcenter.org/wp-content/uploads/Considerations-in-Using-NSC-STHS-Reports.pdf. See footnote 4 below for further discussion of how undocumented students might impact the report's findings.

Identifying Comparison Groups

Which MPS students who did not participate in College Possible could have participated in the program and thus serve as an appropriate comparison group for estimating impacts? This question is central to the evaluation and very difficult to answer. Some comparable students did not apply to the program, which is the first step in participation. Maybe students apply because they are highly motivated, or maybe they are recruited to apply because someone in their lives views them as "at risk." These things are hard to measure but indicate that simply comparing applicants to non-applicants may over- or under-state program effects.

Instead, we turned to data on all MPS students and focused on a group for whom we had a lot of detailed information that could be used to match participants and non-participants. The most critical data are measures of parental educational attainment and the number of family and friends attending college, as these predict college attendance.[2] Unfortunately, these data are available only for students who reached the end of senior year and took the ACT. As a result, we restrict the range of program outcomes that can be examined—for example, most of the students who reached the end of senior year also graduated from high school—but on the other hand the detailed data make it possible to examine impacts on college transitions with more precision.

We used the data and employed a propensity score (PS) matching method known as radius matching. The first step is to model the probability that a high school student participated in College Possible. The estimated value of this probability is known as a "propensity score." The estimated propensity score is then used to match each student in College Possible to comparison students with similar propensity scores (e.g. those within a specified radius distance). In other words, students in the program are compared to other students with similar odds of being in the program, based on measurable characteristics. After matching students in College Possible with similar non-participants, we examine differences in their educational outcomes.[3]

For college admissions activities (number of college admissions applications, number of scholarship applications, and FAFSA completion) the sample is further limited to those survey respondents who indicated plans to attend college immediately following high school, because only students who indicated such plans were asked about admissions outcomes on the Exit Survey. Both this sample and our preferred sample are very well balanced by PS matching on observed

[2] Omitting parent education and the number of family and friends attending college from matching and outcome models results in substantially higher impact estimates for college enrollment, ACT scores, AP/IB courses taken, and high school achievement. Therefore, including these variables is essential for reducing potential biases.
[3] In this analysis, after testing a variety of caliper sizes commonly employed in PS matching (from 0.05 to 0.25 standard deviations) the radius is set to 0.05 standard deviations. This caliper achieves good balance while retaining a large number of College Possible students in the matched analysis sample.

covariates (see the Technical Appendix). Theoretically, constraining the sample further to include only students who applied to College Possible could be an even better way to control for potential bias, because applicants may be more motivated than non-applicants. We tested matching with this subsample as well, but found that PS matching was unsuccessful in balancing on observed covariates. Accepted applicants differ substantially from rejected applicants, severely limiting the number and quality of matches.

For a subset of outcomes, neither of the samples described above is practical. First, for two of the schools in our sample, Riverside University High School and Hmong American Peace Academy, the course taking data provided by MPS displays erratic patterns by grade level. We concluded that the course data from these two schools are unreliable and therefore exclude them when estimating program impacts on the number of AP or IB, math, and science courses taken in grade 12. Second, when estimating impacts on high school graduation, we expand the sample to the full cohort of students who attended any of the eight included MPS high schools in the 2014-15 and 2015-16 school years. This was necessary, because limiting the sample to Exit Survey respondents excludes virtually all non-graduates. Third, college selectivity outcomes such as college type and acceptance rate, are estimated for the college-going sample only. Finally, year-to-year college persistence is estimated using the sample of college-goers from only the first cohort because NSC data on the second cohort's second year in college are not yet available.

Table 1 lists all variables included in the treatment model. This model achieves good balance for the majority of included variables (see Technical Appendix, Tables A-F). To ensure that any remaining imbalances between the participant and non-participant groups do not impact the analyses, we include identical variables in the models used to estimate program impacts (see Estimating Impacts section below).

Table 1: Propensity Score Treatment Model Covariates

School attended in grade 11, indicators for each school:*
• Hamilton High School • Pulaski High School • Riverside University High School • South Division High School • Milwaukee High School of the Arts • Morse Marshall High School • Milwaukee School of Languages • Hmong American Peace Academy
Same school attended grade 10 and 11
Female
Race, indicators for each of the following:*
• White • African American • Hispanic • Asian • American Indian
Eligible for free or reduced price lunch
English Language Learner
Parent education (highest attained either parent), indicators for each of the following:*
• High school or less • Some college • BA or above • Don't know/not applicable
Number of family or friends who attended college
End of Year Cumulative GPA, grade 10
End of Year Cumulative GPA, grade 10, squared
Number of credits earned, grade 10
Attendance rate, grade 10
Number of disciplinary actions, grade 10

** In the interest of completeness, all values of a given variable are listed above. When estimating the propensity score models, one value serves as the reference category and is dropped from calculations.*

These outcome models control for important demographic variables and variables for which matching fails to balance values across the participant and non-participant comparison groups. Including these variables in our outcome models ensures that they do not influence the impact estimates. The following variables are included in all impact estimate models: indicators for each category of race, gender, and English language learner status; free or reduced price lunch status; cumulative GPA in grade 10, as well as a squared version of this term; number of disciplinary actions in the 10th grade; and indicator variables for each school in the sample. Parent educational attainment and number of family and friends who attended college are included in all models except one—the model used to estimate graduation impacts. These variables cannot be included in the high school graduation outcome model, because they are drawn from the Senior Exit Survey, which was completed almost exclusively by graduates. Lastly, for models estimating impacts on course taking in grade 12, the number of AP or IB, math, and science courses taken in grade 11 are included as covariates.

Impacts

Consistent with College Possible's focus, we find evidence that the program positively impacts college enrollment and the selectivity of the colleges that students attend. The results also show that, prior to college enrollment, College Possible has positive impacts on admissions applications, scholarship applications, and FAFSA completion. Moreover, we find some evidence that the program is supporting college preparedness through improved high school graduation rates, ACT Writing and English scores, and increased enrollment in AP/IB, math, and science courses during grade 12. Our analysis does not suggest that College Possible improved intermediate high school outcomes such as grades, attendance, or disciplinary actions.

College Enrollment, Selectivity, and Match

Relative to a comparison group of similar students, students served by College Possible in Milwaukee were 18 percentage points more likely to enroll in college immediately after high school. While roughly three out of five students in the comparison group immediately enrolled in college, four of five College Possible students did so. This impact is not likely due to chance and represents a substantial increase in college participation.[4]

[4] As discussed above, NSC data frequently omits undocumented students. Due to the substantial population of undocumented students in Milwaukee, estimated impacts on college enrollment could be biased if undocumented students are over- or underrepresented in the College Possible population. To test whether the omission of undocumented students in NSC data impacts our findings, we also estimated impacts with a sample omitting Hispanic students. Although undocumented students are not only Hispanic and most Hispanic students are citizens or permanent residents, the majority of undocumented MPS students are Hispanic and neither College Possible nor MPS track undocumented students. For the sample of non-Hispanic students, College Possible students are 14 percentage points more likely to immediately enroll in college relative to the comparison group. Given the similarity of this result to the full-sample impacts described above, it is unlikely that the limitations of NSC data are driving the estimated program impacts presented in Table 2.

Table 2: Impact Estimates: College Enrollment, Persistence, and Selectivity

College Enrollment						
	Mean Difference[1]	Comparison[2]	Participant[3]	p-value		N[4]
Immediate College Enrollment						2005
Mean number	18 pp[5]	62%	80%	0.000	***	
College Persistence (cohort 1 only)	6 pp	76%	82%	0.243		520
College Choice						995
Enrolled in a 4-year institution	18 pp	61%	79%	0.000	***	
Acceptance rate of institution	-6.32 pp	82.50%	76.18%	0.001	***	

Notes: 1. Differences are model estimated average treatment effects for the treated (ATT). 2. Comparison means reported above are the potential outcome means (POMs) estimated by the outcome models. 3. Participant means are the actual mean values for all College Possible participants in the sample. 4. For college enrollment, the sample is restricted to students who completed the MPS Senior Exit Survey and for whom we obtained ACT data. College selectivity measures are estimated with a sample restricted to college-goers. The sample for persistence estimation is further restricted to students in the first sample cohort. 5. Here and throughout the report, "pp" indicates a percentage point difference. 6. For mean differences between the participant group and the comparison group, the key of symbols for statistical significance is as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

One of College Possible's goals is to help students find a better "match" with their college. Often, high-achieving, low-income students attend less-selective colleges with lower-achieving peers, and good matching has positive impacts on both college completion and future earnings.[5] Relative to the comparison group, students served by College Possible were 18 percentage points more likely to attend a 4-year college or university, as shown in Table 2. On average, the institutions College Possible students enrolled in had acceptance rates 6.32 percentage points lower than the institutions that the comparison group attended.

[5] Dillon, E. W. & Smith, J. A. (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics, 35*(1), 45-66. Dillon, E. W. & Smith, J. A. (2017). The consequences of academic match between students and colleges. CESifo Working Paper Series No. 6344. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2933410

Table 3: Comparisons of College Match for College Possible Participants and Non-participants

Students in College Possible									
	ACT Composite			Cumulative GPA			AP/IB Classes		
	Low	Middle	High	Low	Middle	High	Low	Middle	High
Enrolled in a 4-year institution	0.59	0.85	0.88	0.65	0.68	0.93	0.63	0.84	0.90
Acceptance rate of institution	80.40	76.28	72.54	80.34	79.05	72.29	79.20	76.35	73.96
Enrolled in a 4-year institution	0.23	0.47	0.76	0.19	0.54	0.78	0.34	0.52	0.66
Acceptance rate of institution	94.30	87.06	76.26	95.03	85.32	75.35	90.79	84.70	79.60

We also examined college choices by students' high school academic performance. In general, students with better academic records should be expected to attend more selective colleges. Table 3 shows the selectivity indicators, attendance at a 4-year institution and college acceptance rate, by terciles of students' ACT scores, cumulative GPA, and total number of AP and IB courses. Consistent with the impact results in Table 2, Table 3 shows that College Possible students chose more selective institutions, as measured by 4-year attendance and acceptance rates. This is particularly true for the lower and middle terciles of academic performance. For example, College Possible students in the middle tercile of cumulative GPA were 14 percentage points more likely to attend 4-year institutions and attend institutions with lower acceptance rates. This trend is more pronounced for the lowest-achieving tercile. College Possible students in the lowest GPA tercile were 46 percentage points more likely to attend a 4-year college. The more modest difference in acceptance rates for those in the lowest GPA tercile indicates that these College Possible students are choosing to attend less-selective 4-year institutions instead of 2-year colleges. The differences in 4-year attendance and acceptance rates are smaller for the highest achieving students.

It is not clear whether changes in rates of enrollment or match translate into improvement in persistence in college, however. The results are less clear, perhaps because they could be examined for only half of the students in the sample.[6] College Possible students who attended college were six percentage points more likely to continue into their second year, relative to the comparison group. While this difference is meaningful, it may be due to chance.[7]

[6] Persistence was measured using a sample of college-goers only, data for the second cohort are not yet available, resulting in an estimation sample roughly one-quarter as large as the sample used to estimate impacts on college enrollment.

[7] The p-value of 0.243 implies that the probability that the result is different from zero is approximately 24%.

Table 4: Impact Estimates: College Admissions, College Preparation, and High School Performance

College Admissions Activities						
	Mean Difference[1]	Comparison[2]	Participant[3]	p-value		N[4]
College Applications Submitted						1366
Mean number	2.01	3.52	5.52	0.000	***	
At least one	3 pp	95%	98%	0.027	*	
Three or more	33 pp	59%	92%	0.000	***	
Five or more	40 pp	32%	72%	0.000	***	
Scholarship Applications Submitted						
Mean number	1.4	2.36	3.76	0.000	***	
At least one	24 pp	59%	83%	0.000	***	
Three or more	30 pp	40%	70%	0.000	***	
Five or more	17 pp	22%	39%	0.000	***	
FAFSA Completed	11 pp	84%	95%	0.000	***	
College Preparation						
ACT Scores						2005
Composite	0.18	17.17	17.35	0.314		
English	0.58	16.16	16.74	0.019	*	
Math	0.16	17.22	17.38	0.353		
Reading	-0.04	17.59	17.55	0.849		
Science	0.12	17.64	17.76	0.601		
Writing	0.29	6.29	6.58	0.002	**	
College Prep Courses, grade 12						1364
AP or IB	0.29	2.07	2.36	0.034	*	
Math	0.12	1.36	1.48	0.085	+	
Science	0.16	1.11	1.27	0.091	+	

College Preparation						
	Mean Difference[1]	Comparison[2]	Participant[3]	p-value		N[4]
High school graduation, on-time	10 pp	86%	96%	0.000	***	3061
GPA, grd. 11	0	2.77	2.77	0.998		2005
GPA, grd. 12	0	2.78	2.78	0.875		
Attendance rate, grd. 11	0 pp	93%	93%	0.690		
Attendance rate, grd. 12	1 pp	92%	91%	0.592		
Num. disciplinary actions, grd. 11	-0.09	0.34	0.25	0.156		
Num. disciplinary actions, grd. 12	-0.02	0.18	0.16	0.465		

Notes: 1. Differences are model estimated average treatment effects for the treated (ATT). 2. Comparison means reported above are the potential outcome means (POMs) estimated by the outcome models. 3. Participant means are the actual mean values for all College Possible participants in the sample. 4. For mean difference between the participant group and the comparison group, key of symbols for statistical significance as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

College Admissions Activities

One way College Possible helps students access postsecondary education is by assisting with the college application and financial aid processes. When the College Possible program is implemented with fidelity, the data should reflect that participants a) complete more college applications, b) complete more scholarship applications, and c) complete the FAFSA at higher rates. Table 4 shows that there is robust evidence that College Possible increased their students' completion of these three admissions-related activities. Relative to the matched comparison group, on average College Possible students completed approximately 2 more college admissions applications and 1.4 more scholarship applications. The rate at which students completed the FAFSA was 11 percentage points higher than for the matched comparison group. For all of these outcomes, estimated impacts are statistically significant with p-values near zero.

We also find consistent program impacts across several application thresholds. College Possible had large impacts on the percentage of students meeting the three and five application targets for college admissions applications (33 and 40 percentage point gains, respectively). The program also had a small impact on the percentage of students meeting the minimum of one application

completed relative to the comparison group (3 percentage points). The smaller impact at the one application threshold is largely a product of the high rate at which the graduating seniors in this sample completed at least one college application (95%). College Possible students went much further than one application, however. They were much more likely to meet the three and five application thresholds. This greater number of applications is important to College Possible's strategy of getting students beyond the decisions to apply to and enroll in college to also being well matched to the institutions they ultimately attend (see Table 3 above for evidence of matching). Having multiple options may increase students' chances of a good match.

For scholarships, the impact at the threshold of one application is large, with 24 percentage points more participants submitting at least one application relative to the comparison group. There are also sizable impacts at the three (30 percentage points) and five application (17 percentage points) levels.

College Preparation

College Possible prepares students for access to college and postsecondary success by preparing them for college entrance exams and encouraging them to take challenging coursework. To investigate potential impacts on these college preparatory activities, we analyzed whether students served by College Possible had higher ACT scores and enrolled in more challenging courses, relative to the matched comparison group.

As shown in Table 4, there is mixed evidence of College Possible's impact on ACT scores. For the most part, differences in ACT scores between participants and matched non-participants are small and could be due to chance. The only statistically significant improvements in ACT scores for College Possible students were on the English and Writing sections, although these impacts were small. We estimate that, on average, students served by College Possible obtained ACT writing scores 0.29 points higher and ACT English scores 0.58 points higher than comparison group students.[8]

There is more convincing evidence of impacts on students' high school course choices. Relative to the matched comparison group, College Possible students enrolled in 0.29 more AP or IB courses in grade 12. This means that of every three students, approximately one of those enrolled in an additional AP or IB course, relative to the comparison group. This result is both a meaningful improvement in students' college preparation and statistically significant. Impacts on math and science coursework are smaller and both are significant at the 0.10 level only.

[8] These findings echo past research showing small, positive impacts from ACT preparation. See Allensworth, E., Correa, M., & Ponisciak, S. (2008). *From high school to the future: ACT preparation – too much, too late.* Retrieved from https://files.eric.ed.gov/fulltext/ED501457.pdf; Consortium on Chicago School Research at the University of Chicago; Briggs, D. C. (2009). *Preparation for college admission exams.* 2009 National Association for College Admission Counseling (NACAC) Discussion Paper. Retrieved from https://files.eric.ed.gov/fulltext/ED505529.pdf

High School Performance

College Possible may also impact high school performance, although the program does not explicitly aim to influence these kinds of outcomes. Students in College Possible may better understand their college potential and exert more effort in high school. There is evidence that College Possible boosted high school graduation rates. Using an expanded sample of all students who were enrolled in one of the 8 MPS high schools during the 2014-15 and 2015-16 school years, we estimate that College Possible increased students' graduation rates by 10 percentage points (Table 4) relative to the matched comparison group. However, expanding the sample to include non-graduates precluded using parent education or number of family and friends in college either in estimating propensity scores or impacts. These measures could not be included, because they were drawn from Exit Survey data and, by and large, students who did not graduate did not complete the survey. The education levels of parents, other family, and friends likely contributed to this estimated graduation effect and these effects cannot be disentangled from any impact College Possible may have had.

Apart from this possible graduation effect, there is no evidence that College Possible contributed to better performance on general high school performance measures such as GPA, attendance, or number of disciplinary actions. In general, the lack of any measured impact on high school performance measures is not surprising, given that a GPA of 2.0 or higher is required for program admission. Students served by College Possible also have excellent attendance and behavior records prior to program participation. Furthermore, our matching strategy narrows the comparison group to students who were also doing quite well along these measures. It would be difficult to improve on the already high GPAs, excellent attendance, and strong behavioral records for College Possible students.

Heterogeneous Impacts

Participation in College Possible Programming

Some students may benefit from College Possible more than others. To better understand if this is the case, we estimated heterogeneous impacts for several groups. We first investigated impact differences by level of student participation. To do so, we created two groups of College Possible participants – those who attended 50% or more of the program sessions offered (the high participation group), and those who attended less than 50% (the low participation group). We then estimated differences in program impacts between the high and low participation groups. The results of this analysis are presented in Table 5.

Table 5: Impact Estimates by Level of Participation

College Enrollment			
	Differential Impact		p-value
Immediate college enrollment			
Mean Difference	8 pp	*	0.046
College persistence (cohort 1 only)	10 pp		0.153
College Choice			
Enrolled in 4-year institution	23 pp	***	0.000
Acceptance rate of institution	-4 pp	*	0.035
College Admissions Activities			
College Applications Submitted			
Mean number	1.19	***	0.000
At least one	-2 pp		0.341
Three or more	5 pp		0.114
Five or more	25 pp	***	0.000
Scholarship Applications Submitted			
Mean number	1.43	***	0.000
At least one	20 pp	***	0.000
Three or more	23 pp	***	0.000
Five or more	25 pp	***	0.000
FAFSA Completed	3 pp		0.163
College Preparation			
ACT Scores			
Composite	0.07		0.781
English	0.33		0.382
Math	0.04		0.888
Reading	-0.02		0.953
Science	-0.01		0.973
Writing	0.11		0.409

College Preparation			
	Differential Impact		p-value
College Prep Courses, grade 12			
AP or IB	0.22		0.286
Math	0.13		0.222
Science	0.12		0.462
High School Performance			
High school graduation, on-time	5 pp	**	0.009
GPA, grade 11	0.05	***	0.001
GPA, grade 12	0.05	**	0.008
Attendance rate, grade 11	2 pp	***	0.000
Attendance rate, grade 12	3 pp	***	0.001
Num. Disciplinary actions, grade 11	-0.19	*	0.050
Num. Disciplinary actions, grade 12	-0.10		0.108

Note: 1. High vs. low participation is defined as attending 50% or more of program sessions or below 50% respectively. 2. For mean difference between a given participation level and the comparison group, key of symbols for statistical significance as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

With respect to college enrollment, students in the high participation group enrolled immediately in college at a rate 8 percentage points higher and enrolled in 4-year institutions at a rate 23 percentage points higher than their low participation group peers, both statistically significant results. Further, high participation group students enrolled in schools with acceptance rates 3.83 percentage points lower than those attended by students in the lower participation group. The high participation group persisted in college at a much higher rate (10 percentage points) than those in the low participation group, although the results are not statistically significant due to the smaller persistence sample.

The high participation group completed on average 1.19 more college admissions applications than students in the low participation group. College Possible appears to be doing a good job of helping all students reach the goal of three college applications, while higher participation seems to correlate with reaching the five or more application goal – there was a 25 percentage point difference

between the high and low participation groups in the percentages of students who completed 5 college applications or more. This dosage impact was similar for scholarship applications. On average, the high participation group submitted 1.43 scholarship applications more than the low participation group. Additionally, the high participation group hit each of the scholarship application targets at much higher rates than the low participation group. The difference in FAFSA completion rates between the high and low participation groups was more modest (3 percentage points greater for the high participation group) and not statistically significant.

Between the high and low participation groups, differences in ACT scores and college preparatory courses were small and not statistically significant. Nonetheless, we find that high participation group students achieved significantly better outcomes in graduation rates, GPA, and attendance rates, relative to low participation group students. In fact, the analysis detected small but statistically significant negative impacts on attendance and grades for students who had low program attendance rates. It is possible, however, that this is not a product of program participation, but rather a reflection of other challenges that impacted both high school performance and program attendance.

ELL Status

We also estimated heterogeneous impacts by gender, race/ethnicity, free/reduced lunch status, and English-language learner (ELL) status. This allowed us to study whether College Possible had different impacts on students in each of these subgroups. For the most part, estimates of heterogeneous impacts are statistically noisy and do not reflect differential impacts among subgroups. For ELL students, however, there were consistent, measurable differences in College Possible impacts across several outcomes, including immediate college enrollment, college and scholarship applications submitted, FAFSA completion, math and writing ACT scores, and advanced course-taking (Table 6). Differential impact by gender, race/ethnicity, and free/reduced lunch status are included in the Appendix as Tables G, H, and I, respectively.

Table 6: Differential Impacts by ELL Status

College Enrollment			
	Differential Impact		p-value
Immediate College Enrollment			
Mean difference	20 pp	+	0.064
College Persistence (cohort 1 only)	-29 pp		0.178
College Choice			
Enrolled in a 4-year institution	-5 pp		0.696
Acceptance rate of institution	15 pp		0.975
College Admissions Activities			
College Applications Submitted			
Mean number	1.03	*	0.047
At least one	2 pp		0.724
Three or more	25 pp	*	0.026
Five or more	8 pp		0.462
Scholarship Applications Submitted			
Mean number	1.21	*	0.047
At least one	23 pp	*	0.045
Three or more	7 pp		0.582
Five or more	11 pp		0.330
FAFSA Completed	21 pp	*	0.048
College Preparation			
ACT Scores			
Composite	0.85	*	0.047
English	1.12	+	0.064
Math	1.27	**	0.009
Reading	0.42		0.508
Science	0.45		0.512
Writing	1.08	**	0.007

College Preparation			
	Differential Impact		p-value
College Prep Courses, grade 12			
AP or IB	0.79	*	0.014
Math	0.33		0.169
Science	0.41		0.209
High School Performance			
High school graduation, on-time	0 pp		0.984
GPA, grd. 11	0.01		0.849
GPA, grd. 12	-0.01		0.814
Attendance rate, grd. 11	0 pp		0.903
Attendance rate, grd. 12	0 pp		0.720
Num. Disciplinary actions, grd. 11	0.11		0.362
Num. Disciplinary actions, grd. 12	0.06		0.297

Note: 1. Key of symbols for statistical significance as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

Conclusion

Our analysis provides evidence that College Possible Milwaukee promotes immediate college enrollment. College Possible participants at Milwaukee public high schools, when compared to similar non-participating students, were 18 percentage points more likely to enroll in college immediately after graduation. They were also 18 percentage points more likely to attend a 4-year institution and, on average, attended institutions with lower acceptance rates. There is some evidence that College Possible students were more likely to persist into their second year of college, although persistence data is only available for the first cohort and resulting estimates are not statistically significant.

Relative to the matched comparison group, students served by College Possible were much more successful at completing important college admissions activities. Participants on average submitted 2 more college admissions applications and 1.4 more scholarship applications. They were 11 percentage points more likely to have completed a FAFSA. Not only were College Possible students completing more college applications, they were meeting targets of both three and five

college applications at much higher rates than non-participants. This results in students having choices when it comes time to accept an offer and enroll, which may in turn lead to greater success in college if students are better matched to the colleges they choose.

The higher rate of FAFSA completion among students served by College Possible is equally important, because enrollment in college hinges not only on being admitted, but also being able to pay for college. Federal financial aid is indispensable for the low-income students College Possible serves and is only accessible to those who complete the FAFSA. The larger number of scholarship applications completed by the group of participants with high program attendance rates is also an important indicator that College Possible is effectively helping students prepare to finance their college careers.

We also find evidence that College Possible is supporting students toward higher graduation rates, but this impact may be influenced more by parent educational attainment than by College Possible, given that we were unable to control for parent education in the graduation analysis. We find no evidence that College Possible is impacting other high school level outcomes such as GPA, attendance, or disciplinary actions. This is not cause for concern, however, as students in College Possible were already performing well along these dimensions prior to program admittance.

For a number of outcomes there were meaningful and statistically significant differences between students with high program attendance (50% or more) and those who attended less frequently. Relative to College Possible students with lower levels of participation, high participation students were significantly more likely to enroll in college immediately after high school, attend a 4-year institution, and graduate from high school. College admissions activities were greater across the board for students with higher levels of program participation. The high participation group was significantly more likely to meet the program target of 5 college entrance applications or more and to meet all targets for scholarship applications. College Possible students who attended 50% or more of program sessions also experienced significantly improved impacts on immediate college enrollment, enrollment in 4-year colleges, and college persistence. College Possible has an internal benchmark for students to attend at least 50% of program sessions. This evaluation provides substantial support for that benchmark and may encourage the organization to strive to meet that benchmark for more students. Such a goal may call upon College Possible to go the extra mile to support students facing academic or personal struggles during their time in the program.

This is an observational study and as such there is the possibility that our estimates are biased upwards, because program participants have unobserved characteristics that make them different from non-participants. For example, students in College Possible may have more motivation to succeed in school, as displayed by virtue of applying to and participating in the program. We estimate Rosenbaum bounds to address these concerns (see Technical Appendix for details). For

an unobserved characteristic such as motivation to explain the results, highly motivated students would need to be twice as likely to enter College Possible as observationally identical students with lower motivation. In addition, estimated impacts on college enrollment and college admissions activities remain significant after correcting for multiple comparisons and are robust to several propensity score methodologies.

Technical Appendix

In an observational study like this one, many of the characteristics used to screen students for participation in the program are also highly correlated with high school performance and college preparatory outcomes. Propensity Score (PS) Analysis addresses this selection bias. These methods are useful for addressing overt bias and can also address hidden bias, if the estimated propensity scores are correlated with unobserved covariates. The relative strength or weakness of PS methods for a given analysis rests on a number of factors. Most important are the choice of variables to include in the model used to estimate the propensity scores (the treatment model) and the extent of overlap in propensity scores between the treatment and comparison groups. The specific PS method chosen also has some influence on results, though not nearly as much as the previous two.

In this technical appendix, we supplement the information presented in the Methodology section of the main report, going into more detail regarding choice of variables to include in the treatment model, choice of PS method, and covariate balance statistics for our analysis samples.

Specifying the Treatment Model

The equation below represents the log-odds (logit) model we employed for our treatment model. In this equation p represents the propensity score, α is an intercept term, β is a vector of coefficients on the covariates (X) included in the model, and ϵ is an error term.

The first step in our analysis was to test a large number of demographic and pre-program variables (X) for inclusion in the treatment model. High school attendance, discipline, and achievement variables from students' tenth grade year are all good candidates for this model. In some cases, however, multiple measures are highly correlated with each other, diminishing the overall predictive power of the model. For instance, variables for credits attempted and credits earned are highly correlated with each other. We chose to exclude credits attempted from our model. Another example was correlation between the total number of disciplinary actions in the year and another variable counting the number of suspension days. Here we prefer inclusion of number of disciplinary actions over number of suspension days, because very few students experienced suspension.

Other variables that are logical candidates for the treatment model could not be included, because they perfectly predict selection into either the treatment or control condition. For example, we tested inclusion of a variable that counted the number of gaps in a student's enrollment for the tenth grade year. However, no students with gaps in their tenth grade enrollment record were admitted to College Possible. In this case, we opt to use enrollment gaps in the ninth grade year instead.

Finally, in specifying the treatment model, there is also the question of whether interaction terms or higher order terms should be included. We tested various interactions of demographic characteristics and high school record data; as well as squared terms for GPA, attendance rate, and number of disciplinary records in the tenth grade year. Including a squared GPA term improves the model, while the other squared terms do not. None of the interaction terms tested improves prediction.

The final list of covariates included in the propensity score treatment model is found in Table 1 in the Methodology section of this report.

Overlap

In our analyses, there is good overlap for nearly all students in the treatment condition (participation in College Possible). There are many non-participants, however, for whom there are no good matches in the treatment group. We address this challenge, in part by limiting our sample for analysis to the sub-sample of Senior Exit Survey respondents who indicated plans to attend college immediately following high school (see Methodology section). In addition, for the impacts presented in this report, we limit our discussion to estimated impacts for participants (Average Treatment Effect for the Treated) rather than for the full sample (Average Treatment Effect).

Selecting a Propensity Score Method

We tested a variety of PS matching algorithms as well as PS weighting using the above treatment model. The goal of these tests was to identify the PS method that would achieve the best balance on covariates included in the treatment model. We also aimed to identify a model that would retain as many observations as possible for estimating impacts. We tested the following Propensity Score Matching algorithms: 1) nearest neighbor matching with one, two, and three matched comparisons per treated observation; 2) radius matching with various caliper sizes; 3) kernel matching with varying bandwidths, as well as with and without trimming outliers in the propensity score distribution. We also tested inverse propensity score weighting. Similar levels of balance were achieved using matching to the three nearest neighbors, radius matching with caliper size 0.05 standard deviations, kernel matching with 0.80 bandwidth and no trimming, and inverse probability weighting.

We ultimately decided on propensity score matching using the radius method with a caliper of 0.05 standard deviations. This method achieves the best balance on covariates of all methods tested. We then estimated program impacts using linear regression models with the matched sample, including as covariates all demographic variables and school indicators, as well as other pre-program variables for which the matched sample was still somewhat unbalanced. These regression adjusted impact estimates correct for residual bias after matching. Results may continue to be biased, however, if both models are incorrectly specified. For instance, in our analysis of high school graduation, our results are only tentative, because parent education is an important predictor of academic performance. Excluding this information likely means that both our treatment and outcome models are incorrectly specified.

Matched Sample and Covariate Balance

The sample for estimating treatment effects could be as broad as the sample of all MPS students for whom we obtained administrative data or as narrow as the sample of students who applied to College Possible. Our preferred sample consists of those students who completed the MPS administered Senior Exit Survey and for whom we obtained ACT data from MPS. This survey-based sample, though it excludes some program participants, is preferable because it allows us to control for parents' educational attainment and number of family and friends who attended college, both of which are strong predictors of college enrollment. This sample still retains a large number of program participants (N=428), including those who, in spite of their college intentions when applying to College Possible, did not plan to proceed directly from high school to college when asked at the end of their senior year. Limiting the sample to students with ACT scores allows us to use a consistent sample for ACT and high school performance outcomes. We find that limiting to survey respondents with ACT scores produces no substantive change in impact estimates for the high school performance outcomes compared estimates from the larger sample of all survey respondents.

Table A presents a summary of the balance achieved when using PS radius matching with a caliper of 0.05 standard deviations for our main analytic sample of Exit Survey respondents for whom we have ACT score data. The table includes mean values for program participant and comparison groups after matching for each variable included in the treatment model. It also reports p-values from student's t-tests, indicating the statistical significance of differences between the group means. Finally, it includes the percent bias remaining after matching. As you can see below, matching achieves very good balance for covariates included in the treatment model overall. The majority of percent bias values are below 5 percent. Where percent bias is at or above 5 percent, we included those covariates in the regression adjusted outcome models, thereby controlling for the remaining imbalance. Balance statistics for the alternate samples used to estimate college

admissions activities, course taking behavior, and high school graduation outcomes follow in Tables A through F.

Also note that balance reported below is for estimates of the average treatment effect on the treated. That is, matching was used only to identify a comparison group for students served by College Possible. We were unable to achieve sufficient balance on covariates when we attempted to also identify matches for each comparison observation from among the participant group.

Table A: Covariate Balance Summary—Radius Matching, Caliper=0.05 Standard Deviations Exit Survey Respondent with ACT Data Sample

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School, grade 11				
Hamilton High School	14%	13%	0.89	-0.9
Pulaski High School	10%	11%	0.795	1.6
Riverside University High School	17%	16%	0.627	-3.3
South Division High School	12%	13%	0.848	1.3
Milwaukee High School of the Arts	15%	16%	0.921	0.7
Morse Marshall High School	13%	14%	0.762	2.2
Milwaukee School of Languages	8%	8%	0.977	-0.2
Hmong American Peace Academy	9%	9%	0.875	-1.4
Same school, grade 10 and 11	97%	98%	0.786	1.1
Female	64%	66%	0.519	4.3
Race				
White	9%	9%	0.846	-1.2
African American	46%	47%	0.778	1.9
Hispanic	23%	22%	0.62	-3.3
Asian	22%	22%	0.811	1.9
American Indian	0%	0%	0.707	1.9
Free/reduced lunch	89%	90%	0.778	1.5
English Language Learner	9%	9%	0.828	-1.4

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
Parent education (highest attained either parent)				
High school or less	45%	45%	0.906	0.8
Some college	27%	26%	0.905	-0.8
BA or above	13%	14%	0.969	0.2
Don't know/not applicable	15%	15%	0.957	-0.4
Num. family/friends who attended college	2.46	2.45	0.782	-1.8
GPA, grade 10	2.82	2.83	0.726	1.8
GPA, grade 10, squared	8.23	8.32	0.683	2.5
Num. credits earned, grade 10	6.62	6.63	0.836	0.8
Attendance rate, grade 10	95%	95%	0.851	0.6
Num. disciplinary actions, grade 10	0.26	0.26	0.935	-0.2
Sample Size	**1577**	**428**		

Table B: Covariate Balance Summary—Radius Matching, Caliper=0.05 Standard Deviations Exit Survey Respondent with ACT Data Sample—Schools with Reliable Course Data Only

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School, grade 11				
Hamilton High School	19%	17%	0.686	-3.2
Pulaski High School	14%	14%	0.924	0.7
Riverside University High School	-	-	-	-
South Division High School	18%	17%	0.896	-1.2
Milwaukee High School of the Arts	21%	20%	0.833	-1.9
Morse Marshall High School	18%	20%	0.474	6.8
Milwaukee School of Languages	11%	11%	0.941	-0.7

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
Hmong American Peace Academy	-	-	-	-
Same school, grade 10 and 11	98%	98%	0.76	1.3
Female	66%	68%	0.484	5.3
Race				
White	12%	11%	0.756	-2.5
African American	46%	49%	0.577	4.3
Hispanic	28%	26%	0.445	-6.2
Asian	13%	14%	0.707	2.9
American Indian	0%	1%	0.658	2.9
Free/reduced lunch	90%	91%	0.765	1.7
English Language Learner	10%	10%	0.998	0
Parent education (highest attained either parent)				
High school or less	47%	46%	0.772	-2.3
Some college	28%	29%	0.71	2.9
BA or above	13%	12%	0.952	-0.4
Don't know/not applicable	13%	13%	0.993	-0.1
Num. family/friends who attended college	2.38	2.36	0.764	-2.1
GPA, grade 10	2.80	2.81	0.732	2
GPA, grade 10, squared	8.11	8.22	0.657	3.1
Num. credits earned, grade 10	6.69	6.69	0.997	0
Attendance rate, grade 10	94%	95%	0.744	1.2
Num. disciplinary actions, grade 10	0.24	0.25	0.928	0.2
Sample Size	**1136**	**329**		

Note: 1. Data on course taking from Riverside University High School and Hmong American Peace Academy displayed erratic patterns across school years. Data from these schools were therefore excluded from analysis of course taking outcomes.

Table C: Covariate Balance Summary—Radius Matching, Caliper=0.10 Standard Deviations College Bound Exit Survey Respondent Sample

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School, grade 11				
Hamilton High School	13%	13%	0.933	-0.6
Pulaski High School	10%	10%	0.899	0.8
Riverside University High School	19%	18%	0.696	-3
South Division High School	11%	11%	0.866	1.2
Milwaukee High School of the Arts	18%	17%	0.886	-1.2
Morse Marshall High School	11%	12%	0.748	2.4
Milwaukee School of Languages	8%	9%	0.771	2.4
Hmong American Peace Academy	10%	10%	0.892	-1.4
Same school, grade 10 and 11	97%	97%	0.663	2
Female	67%	68%	0.815	1.7
Race				
White	9%	9%	0.959	-0.4
African American	48%	48%	0.892	1
Hispanic	21%	20%	0.764	-2.1
Asian	22%	23%	0.912	1
American Indian	0%	0%	0.629	1.9
Free/reduced lunch	89%	89%	0.857	1.1
English Language Learner	8%	8%	0.925	-0.6
Parent education (highest attained either parent)				
High school or less	44%	43%	0.851	-1.4
Some college	27%	28%	0.845	1.5
BA or above	14%	15%	0.734	2.3
Don't know/not applicable	15%	14%	0.751	-2.5

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
Num. family/friends who attended college	2.48	2.48	0.999	0
GPA, grade 10	2.85	2.87	0.713	2.1
GPA, grade 10, squared	8.43	8.53	0.677	2.9
Num. credits earned, grade 10	6.61	6.62	0.831	0.9
Attendance rate, grade 10	95%	95%	0.947	0.2
Num. disciplinary actions, grade 10	0.28	0.26	0.754	-0.9
Sample Size	**1010**	**356**		

Note: College admissions related outcomes were derived from MPS Senior Exit Survey items completed only by students who indicated plans to attend college immediately following college. For these outcomes, therefore, matching was conducted only among participants and non-participants who indicated college plans.

Table D: Covariate Balance Summary— Radius Matching, Caliper=0.05 Standard Deviations Sample of All in Cohort Enrolled in One of the Eight MPS High Schools in their Junior Year

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School, grade 11				
Hamilton High School	13%	13%	0.933	-0.6
Pulaski High School	10%	10%	0.899	0.8
Riverside University High School	19%	18%	0.696	-3
South Division High School	11%	11%	0.866	1.2
Milwaukee High School of the Arts	18%	17%	0.886	-1.2
Morse Marshall High School	11%	12%	0.748	2.4
Milwaukee School of Languages	8%	9%	0.771	2.4
Hmong American Peace Academy	10%	10%	0.892	-1.4
Same school, grade 10 and 11	97%	97%	0.663	2

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School, grade 11				
Hamilton High School	14%	13%	0.771	-1.7
Pulaski High School	12%	13%	0.741	2.1
Riverside University High School	17%	16%	0.621	-3.1
South Division High School	12%	12%	0.965	-0.3
Milwaukee High School of the Arts	16%	16%	0.798	-1.8
Morse Marshall High School	13%	15%	0.466	5
Milwaukee School of Languages	8%	8%	0.925	-0.6
Hmong American Peace Academy	7%	8%	0.835	1.6
Same school, grade 10 and 11	98%	98%	0.907	0.4
Female	64%	65%	0.794	1.7
Race				
White	10%	9%	0.722	-2.2
African American	47%	48%	0.738	2.2
Hispanic	23%	22%	0.631	-3.1
Asian	20%	20%	0.751	2.3
American Indian	1%	1%	0.876	0.9
Free/reduced lunch	89%	89%	0.668	2.2
English Language Learner	10%	9%	0.886	-0.9
GPA, grade 10	2.79	2.81	0.611	2.5
GPA, grade 10, squared	8.10	8.19	0.665	2.6
Num. credits earned, grade 10	6.59	6.62	0.560	2.2
Attendance rate, grade 10	94%	95%	0.432	2.5
Num. disciplinary actions, grade 10	0.34	0.31	0.691	-1.3
Sample Size	**2585**	**476**		

Note: Variables for parent education level and number of family friends who attended college were excluded from matching for this model used to estimate impacts for the outcome high school graduation, because these variables were derived from Senior Exit Survey data.

Table E: Covariate Balance Summary— Radius Matching, Caliper=0.05 Standard Deviations Sample of Exit Survey Respondents with ACT Data Sample who Attended College Immediately After High School

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School - grade 11				
Hamilton High School	12%	11%	0.784	-2.2
Pulaski High School	9%	10%	0.647	3.9
Riverside High School	21%	18%	0.352	-9.1
South Division High School	9%	11%	0.388	7.6
Milwaukee High School of the Arts	13%	15%	0.621	4.8
Morse Marshall High School	11%	15%	0.174	14
Milwaukee School of Languages	9%	9%	0.928	-1
Hmong American Peace Academy	15%	10%	0.073	-25.3
Same school, grade 10 and 11	98%	98%	0.715	1.9
Female	68%	67%	0.810	-2.2
Race				
White	8%	8%	0.885	1.2
African American	44%	48%	0.433	7.5
Hispanic	19%	21%	0.517	5.9
Asian	28%	22%	0.124	-18
American Indian	1%	0%	0.558	-4.1
Free/reduced lunch	87%	90%	0.415	6.2
English Language Learner	7%	8%	0.506	5.8
Parent education (highest attained either parent)				
High school or less	42%	42%	0.881	1.5
Some college	24%	29%	0.270	10.5
BA or above	13%	15%	0.714	3.2
Don't know/not applicable	21%	14%	0.60	-19.8

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
Num. family/friends who attended college	2.51	2.47	0.713	-3.5
GPA, grade 10	2.90	2.90	0.925	0.7
GPA, grade 10, squared	8.70	8.72	0.938	0.7
Num. credits earned, grade 10	6.49	6.65	0.026	12
Attendance rate, grade 10	96%	95%	0.760	-1.3
Num. disciplinary actions, grade 10	0.22	0.24	0.789	0.9
Sample size	**911**	**363**		

Table F: Covariate Balance Summary— Radius Matching, Caliper=0.05 Standard Deviations Sample of Cohort One Exit Survey Respondents with ACT Data Sample who Attended College Immediately After High School

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
School - grade 11				
Hamilton High School	12%	14%	0.73	3.5
Pulaski High School	9%	6%	0.384	-7.8
Riverside High School	21%	20%	0.815	-2.8
South Division High School	9%	11%	0.507	6.7
Milwaukee High School of the Arts	13%	12%	0.774	-3.2
Morse Marshall High School	11%	14%	0.469	8.3
Milwaukee School of Languages	9%	10%	0.912	1.4
Hmong American Peace Academy	15%	14%	0.65	-8.4
Same school, grade 10 and 11	98%	98%	0.77	1.8
Female	68%	68%	0.929	1

Balance Statistics	Mean		t-test	% bias
	Comparison	Participant	p-value	
Race				
White	8%	8%	0.864	1.7
African American	45%	45%	0.909	1.3
Hispanic	19%	19%	0.991	-0.1
Asian	28%	27%	0.82	-3.3
American Indian	1%	1%	0.977	0.3
Free/reduced lunch	87%	88%	0.761	2.9
English Language Learner	7%	7%	0.844	1.9
Parent education (highest attained either parent)				
High school or less	42%	43%	0.776	3.3
Some college	24%	23%	0.89	-1.5
BA or above	13%	15%	0.727	3.6
Don't know/not applicable	21%	19%	0.61	-6.9
Num. family/friends who attended college	2.51	2.50	0.928	-1.1
GPA, grade 10	2.89	2.90	0.906	1
GPA, grade 10, squared	8.68	8.74	0.867	1.8
Num. credits earned, grade 10	6.49	6.51	0.828	1.3
Attendance rate, grade 10	96%	96%	0.878	0.7
Num. disciplinary actions, grade 10	0.22	0.22	0.987	-0.1
Sample size	**365**	**155**		

Heterogeneous Impacts by Gender, Race/Ethnicity, and Free/Reduced Lunch Status

Some students may benefit from College Possible more than others. To better understand if this is the case, we estimated heterogeneous impacts for several groups. Heterogeneous impacts by College Possible participation rates and ELL status are shown above. Tables G-I below display heterogeneous results by gender, race/ethnicity, and free/reduced price lunch recipiency.

Table G: Differential Impact by Gender

College Enrollment			
	Differential Impact		p-value
Immediate College Enrollment			
Mean difference	-3.56 pp		0.526
College Persistence (cohort 1 only)	-8.75 pp		0.419
College Choice			
Enrolled in a 4-year institution	3.99 pp		0.553
Acceptance rate of institution	1.88 pp		0.444
College Admissions Activities			
College Applications Submitted			
Mean number	0.17		0.630
At least one	5.40 pp		0.404
Three or more	3.42 pp		0.626
Five or more	3.95 pp		0.548
FAFSA Completed	0.76 pp		0.858
College Preparation			
ACT Scores			
Composite	-0.34		0.360
English	-0.19		0.715
Math	-0.26		0.494
Reading	-0.61		0.224
Science	-0.26		0.580
Writing	-0.06		0.779

College Preparation			
	Differential Impact		p-value
College Prep Courses, grade 12			
AP or IB	0.97	***	0.001
Math	0.43	**	0.006
Science	0.13		0.522
High School Performance			
High school graduation, on-time	4.70 pp		0.113
GPA, grd. 11	0		0.845
GPA, grd. 12	-0.02		0.354
Attendance rate, grd. 11	-0.43 pp		0.578
Attendance rate, grd. 12	0.28 pp		0.752
Num. Disciplinary actions, grd. 11	-0.14		0.296
Num. Disciplinary actions, grd. 12	-0.05		0.509

Note: 1. This table shows impacts on female versus male. The indicator was define with a value of one for female and zero for male. 2. For differential impact, key of symbols for statistical significance as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

Table H: Differential Impact by Race/Ethnicity

College Enrollment									
	African-American			Hispanic			Asian		
	Differential Impact		p-value	Differential Impact		p-value	Differential Impact		p-value
Immediate College Enrollment									
Mean difference	-0.22 pp		0.982	9.26 pp		0.385	-10.23 pp		0.316
College Persistence (cohort 1 only)	-18.18 pp		0.251	-3.47 pp		0.842	-3.14 pp		0.867
College Choice									
Enrolled in a 4-year institution	2.55 pp		0.792	21.25 pp	+	0.071	-4.81 pp		0.635
Acceptance rate of institution	1.61 pp		0.718	-3.67 pp		0.435	5.27 pp		0.265
College Admissions Activities									
	African-American			Hispanic			Asian		
	Differential Impact		p-value	Differential Impact		p-value	Differential Impact		p-value
College Applications Submitted									
Mean number	0.45		0.453	0.25		0.702	-0.52		0.418
At least one	0.26 pp		0.920	2.84 pp		0.451	-2.12 pp		0.367
Three or more	8.63 pp		0.377	25.14 pp	*	0.021	10.04 pp		0.354
Five or more	10.64 pp		0.361	6.44 pp		0.619	5.10 pp		0.689
Scholarship Applications Submitted									
Mean number	0.69		0.275	0.99		0.138	0.81		0.223
At least one	8.46 pp		0.412	12.12 pp		0.297	11.80 pp		0.265
Three or more	19.29 pp	+	0.086	25.67 pp	*	0.038	20.25 pp	+	0.092
Five or more	9.75 pp		0.392	8.90 pp		0.468	7.84 pp		0.541
FAFSA Completed	4.48 pp		0.380	5.88 pp		0.425	3.22 pp		0.609

College Preparation									
	African-American			Hispanic			Asian		
	Differential Impact		p-value	Differential Impact		p-value	Differential Impact		p-value
Composite	-0.14		0.833	0.25		0.719	-1.07		0.123
English	-1.23		0.205	-0.35		0.741	-2.49	*	0.016
Math	-0.35		0.548	-0.45		0.482	-0.71		0.292
Reading	0.15		0.867	0.58		0.557	-1.29		0.190
Science	0.86		0.307	1.21		0.185	0.28		0.763
Writing	0.31		0.278	0.79	*	0.016	0.34		0.277
College Prep Courses, grade 12									
AP or IB	-0.30		0.517	-0.43		0.386	-0.08		0.878
Math	-0.12		0.630	-0.13		0.632	0.12		0.687
Science	0.60	*	0.016	0.22		0.417	0.64	+	0.083
High School Performance									
	African-American			Hispanic			Asian		
	Differential Impact		p-value	Differential Impact		p-value	Differential Impact		p-value
GPA, grd. 11	0		0.931	0.03		0.470	-0.01		0.722
GPA, grd. 12	0		0.935	0.04		0.516	-0.04		0.472
Attendance rate, grd. 11	1.33 pp		0.304	2.09 pp		0.171	1.41 pp		0.287
Attendance rate, grd. 12	0.99 pp		0.547	0.98 pp		0.593	1.14 pp		0.511
Num. Disciplinary actions, grd. 11	-0.11		0.456	-0.21		0.265	0.03		0.781
Num. Disciplinary actions, grd. 12	0.04		0.670	0.04		0.645	0.10		0.116

Note: 1. For differential impact, key of symbols for statistical significance as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

Table I: Differential Impact by Free/Reduced Lunch Status

College Enrollment			
	Differential Impact		p-value
Immediate College Enrollment			
Mean difference	9.47 pp		0.182
College Persistence (cohort 1 only)	3.87 pp		0.722
College Choice			
Enrolled in a 4-year institution	14.58 pp	+	0.094
Acceptance rate of institution	-8.86 pp	**	0.002
College Admissions Activities			
College Applications Submitted			
Mean number	-0.06		0.897
At least one	2.78 pp		0.425
Three or more	5.95 pp		0.399
Five or more	3.52 pp		0.711
Scholarship Applications Submitted			
Mean number	-0.08		0.864
At least one	3.33 pp		0.684
Three or more	-0.14 pp		0.988
Five or more	1.17 pp		0.905
FAFSA Completed	3.18 pp		0.443
College Preparation			
ACT Scores			
Composite	0.73		0.129
English	0.44		0.518
Math	0.15		0.754
Reading	0.91		0.206
Science	1.01		0.115
Writing	0.31		0.206

College Preparation			
	Differential Impact		p-value
College Prep Courses, grade 12			
AP or IB	0.81	*	0.049
Math	0		0.994
Science	0.33		0.157
High School Performance			
High school graduation, on-time	2.74 pp		0.496
GPA, grd. 11	0.01		0.640
GPA, grd. 12	0.01		0.873
Attendance rate, grd. 11	-0.46 pp		0.554
Attendance rate, grd. 12	-0.58 pp		0.676
Num. Disciplinary actions, grd. 11	0.08		0.540
Num. Disciplinary actions, grd. 12	0.07		0.311

Note: 1. For differential impact, key of symbols for statistical significance as follows: *** 0.001 level, ** 0.01 level, * 0.05 level, + 0.10 level.

Sensitivity and Robustness Analyses

As a check on the reliability of the findings presented above, we present the findings of multiple sensitivity and robustness analyses. These include the Benjamini-Hochberg procedure to control for multiple comparisons, testing several estimation strategies to ensure that results are not driven by the choice of estimation, and calculating Rosenbaum bounds to test the sensitivity of the results to an unobserved confounding variable.

Multiple Comparisons Analysis

Estimating impacts for many outcomes can produce false discoveries by chance. Choosing a significance level of 0.05 implies that 5% of statistically significant estimates will be caused by random chance, not because of actual program impacts. The Benjamini-Hochberg procedure is a commonly accepted method of correcting for the multiple comparisons problem by accounting for the total number of statistical tests as well as the strength of the estimates, as measured by p-values.

Table J below shows the results of the Benjamini-Hochberg procedure for the main impacts of College Possible shown above in Tables 2 and 4.The impact estimates are ranked in ascending order by p-values. We then calculate a critical value equal to the rank multiplied by a false discovery rate (chosen to be 0.05), divided by the total number of tests (in this case, 29). If an estimate's p-value is less than the critical value, then the result is considered statistically significant.

As Table J shows, after correcting for multiple comparisons, the key impact estimates remain statistically significant. College Possible's impacts on college enrollment, enrollment in a 4-year institution, high school graduation, college and scholarship applications, and FAFSA completion all remain significant. Impacts on completing at least one college application and AP/IB course taking, which were statistically significant using the typical 0.05 threshold, are no longer significant when controlling for multiple comparisons.

Table J: Results of the Benjamini-Hochberg Procedure for Multiple Comparisons

Outcome	p-value	Rank	B-H Critical Value	Statistically Significant at 0.05 false discovery rate
Immediate college enrollment	0.000	1	0.002	Yes
Enrolled in 4-year institution	0.000	2	0.003	Yes
College applications - mean number	0.000	3	0.005	Yes
College applications - three or more	0.000	4	0.007	Yes
College applications - five or more	0.000	5	0.009	Yes
Scholarship applications - mean number	0.000	6	0.010	Yes
Scholarship applications - at least one	0.000	7	0.012	Yes
Scholarship applications - 3 or more	0.000	8	0.014	Yes
Scholarship applications - 5 or more	0.000	9	0.016	Yes
FAFSA completed	0.000	10	0.017	Yes
Graduated high school	0.000	11	0.019	Yes
Acceptance rate of institution	0.000	12	0.021	Yes
ACT Writing	0.002	13	0.022	Yes
ACT English	0.019	14	0.024	Yes
College applications - at least one	0.027	15	0.026	No

AP or IB courses, grade 12	0.034	16	0.028	No
Math courses, grade 12	0.085	17	0.029	No
Science courses, grade 12	0.091	18	0.031	No
Num. disciplinary actions, grade 11	0.156	19	0.033	No
College persistence (cohort 1 only)	0.243	20	0.034	No
ACT Composite	0.314	21	0.036	No
ACT Math	0.353	22	0.038	No
Num. disciplinary actions, grade 12	0.465	23	0.040	No
Attendance rate, grade 12	0.592	24	0.041	No
ACT Science	0.601	25	0.043	No
Attendance rate, grade 11	0.690	26	0.045	No
ACT Reading	0.849	27	0.047	No
GPA, grade 12	0.875	28	0.048	No
GPA, grade 11	0.998	29	0.052	No

Robustness to Alternative Estimation Strategies

We estimated impacts using a number of alternative methods. First we employed two alternate PS analysis approaches-- PS nearest-neighbor matching and PS weighting. Second, we compared our regression adjusted results to impact estimates based on simple mean comparisons for the participant and non-participant matched samples. These alternative analyses support the findings presented above. While mean values vary slightly by method, the pattern of statistically significant impacts and the qualitative conclusions we draw from them remain unchanged.

Sensitivity Analysis

Propensity score matching can only imperfectly correct for the selection bias inherent in an observational analysis like this one. While matching can effectively remove biases due to observed characteristics (e.g. previous academic preparation), it does not necessarily correct for unobserved characteristics (e.g. student or family motivation). While there is no way of knowing how much unobserved characteristics may impact the estimates above, we can test the sensitivity of the results to a hypothetical unobserved variable. Rosenbaum bounds test this sensitivity by re-calculating p-values while accounting for an unobserved, binary variable that is correlated with both College Possible attendance and the tested outcome.

Table K shows the Rosenbaum bounds for the primary impact estimates that remained statistically significant after controlling for multiple comparisons (see Table J). For each outcome, several Rosenbaum bounds are calculated for varying strengths of the selection bias caused by an unobserved characteristic. The strength reflects the difference in probability that a student with the confounding characteristic enrolls in College Possible, relative to a student with identical observable characteristics who does not share the confounding characteristic. For example, if the strength of the selection bias is 2.00, the student with the confounding characteristic would be twice as likely to enroll in College Possible relative to an otherwise identical student without the confounding characteristic.

The results in Table K show that immediate college enrollment, on-time high school graduation, completing 3 or more college applications or scholarship applicants, and FAFSA completion are robust to the hypothetical confounder until the selection bias is greater than 2.00. These results imply that selection bias would need to be very strong to account for the estimated impacts.

Table K: Sensitivity of p-values to an Unobserved, Confounding Variable

Strength of Selection Bias							
Outcome	**1.00**	**1.25**	**1.50**	**1.75**	**2.00**	**2.25**	**2.50**
Immediate college enrollment	0.000	0.000	0.000	0.000	0.017	0.573	0.988
Enrolled in a 4-year institution	0.000	0.000	0.000	0.084	0.725	0.990	1.000
On-time high school graduation	0.000	0.000	0.000	0.000	0.011	0.399	0.941
College applications submitted - 3 or more	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Scholarship applications submitted - 3 or more	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FAFSA completed	0.000	0.000	0.000	0.001	0.041	0.260	0.632

Wisconsin Hope Lab Mission

The Wisconsin HOPE Lab was established in 2013 by Dr. Sara Goldrick-Rab on the University of Wisconsin–Madison campus to engage in translational research aimed at improving equitable outcomes in postsecondary education. The Lab is housed in the School of Education and is led by Acting Director Dr. Jed Richardson. For more information, see www.wihopelab.com.

Acknowledgments

The authors would like to thank the Milwaukee Public Schools for providing data used in this report. The authors would also like to thank Janet Trembley and Alison Bowman, who designed this report.

University of Wisconsin–Madison | L139 Education | 1000 Bascom Mall | Madison, Wisconsin 53706 | wihopelab@wcer.wisc.edu

The Great Lakes Higher Education Guaranty Corporation, through a gift to the Wisconsin HOPE Lab, provided financial support for this work.

The Wisconsin HOPE Lab is supported by the Wisconsin Center for Education Research at the School of Education, University of Wisconsin-Madison. Funding by Great Lakes.

wihopelab.com