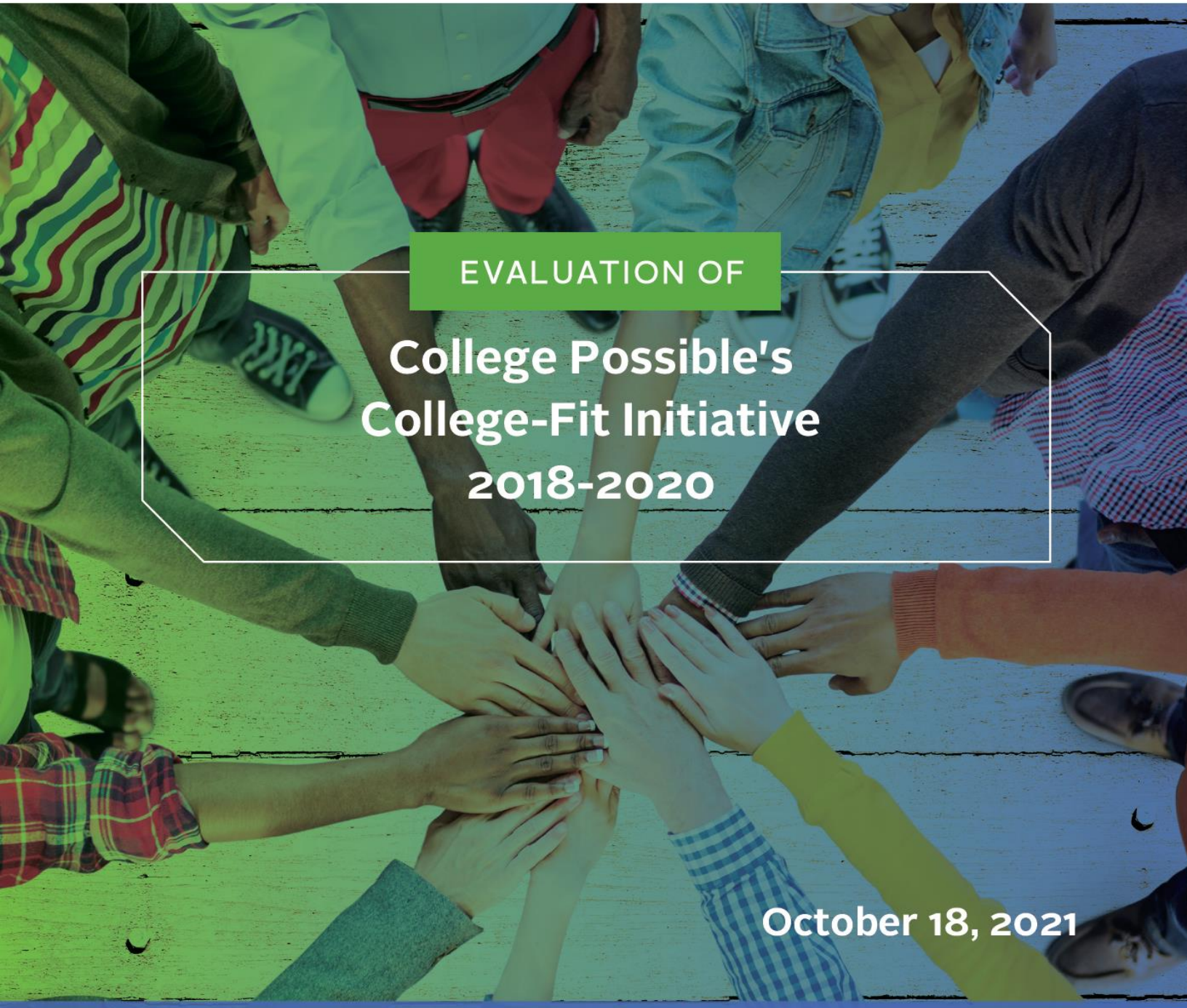
COLLEGE POSSIBLE™



EVALUATION OF

College Possible's College-Fit Initiative 2018-2020

October 18, 2021

Submitted to

College Possible
755 Prior Ave N, Unit 210,
Saint Paul, MN 55104,

Submitted by

Westat, Inc.
An Employee-Owned Research Corporation®
1600 Research Blvd,
Rockville, MD 20850

Evaluation of College Possible's College-Fit Initiative

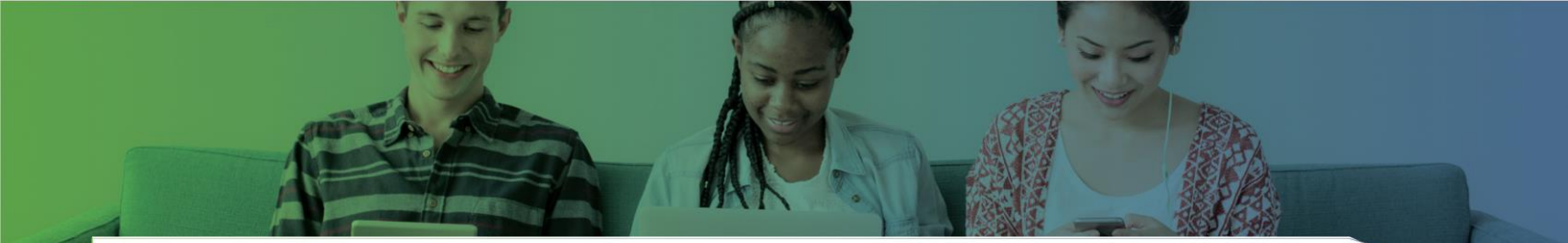
Theory of Change

College Possible is a national nonprofit organization that makes college admission and success more likely for low-income students through an intensive curriculum of coaching and support. College Possible utilizes AmeriCorps members as coaches to provide near-peer coaching to students. By harnessing the energy and idealism of recent college graduates as near-peer coaches to groups of students working towards a shared goal, the program hopes to build a culture of trust and high expectations in which students excel. Coaches provide (1) academic support that begins with ACT/SAT test preparation and continues with, (2) college application assistance, (3) financial aid consulting (including financial literacy, scholarship search & renewal, and strengthening financial fitness), (4) guidance in the college transition and (5) ongoing, targeted, and personalized support toward college degree completion.

College Possible has claimed success in keeping low-income students engaged in school, by working to ensure they graduate high school on time and earn admission to college. College Possible prioritizes the continuous improvement of their services to promote every student served to go as far as their talent, skills, and motivation might take them. One area of programming with an opportunity to intensify their supports is *college fit and selection*. It is claimed that College Possible students enroll in college at a rate that far exceeds the national average for low-income students. However, improving the process by which students select their best-fit college would ideally increase their odds of persisting and graduating – a goal that benefits the student as well as the future workforce needs of the nation.

Context

Although the topic of college choice is woven extensively into College Possible's curriculum, many dimensions of "fit" may be poorly contextualized by their students. College Possible attributes this, in part, to recent research that indicates adolescents have a limited ability to make well-informed, long-term decisions because their executive functioning is still maturing. Ben Castleman, a leading researcher in college access and success for low-income students, has noted that high-school-aged students, in general, are less capable of undertaking the type of careful reasoning that is needed to get to and through college. This deficit in long-term decision-making is further compounded when low-income students are faced with trying to meet needs that high-income students simply do not face. Students from low-income families often struggle to find colleges that are a good match for their abilities *because* their decisions are driven almost exclusively by affordability. College Possible concluded a two-year evaluation in 2018, funded by the U.S. Department of the Treasury, showing this trend. In this study, they asked first-year college students to select all of the reasons why they decided to attend their selected college; 85% cited financial considerations (i.e., cost of attendance and financial aid provided) as the determining factor for their college decision. Conversely, only 35-40% of students indicated that institutional factors such as being close to home, location, size, or the types of academic programs offered also influenced their decisions.



Evaluation of College Possible's College-Fit Initiative

When a student considers only the financial implications of college attendance, this often results in an "undermatch." This is because students choose a seemingly safer financial decision of a nonselective four-year college (which often have low graduation rates) or a two-year college (which traditionally has low degree completion and transfer rates). These factors decrease the likelihood that a student will complete their degree, which can be financially devastating for a low-income student. They do not have the financial security a college degree can offer and at the same time have accumulated additional debt with no added earning power to repay it. When students select and enroll in colleges that are a good fit for their interests, abilities, and circumstances, they are far more likely to persist and earn their college degrees. To address these issues College Possible added two new supports intended to better position their students to find and enroll in their best-fit college. First, they sought to improve the quality of information available for consideration as students decide where to apply and ultimately enroll. Second, they sought to engage families and trusted adults in this process in more meaningful ways. By combining these two supports, College Possible believes they can improve students' ability to select a college where students have the greatest likelihood of persisting and earning a college degree.

Areas of Exploration

To address concerns that students are not effectively considering essential information related to college-fit, College Possible made two program changes. First, College Possible created an advanced scoring system for college prioritization, review, and selection that incorporates internal College Possible data, externally available college data, and student preferences. This support would in theory allow students and their coaches to take a more sophisticated approach to developing target college lists and comparing college options. By offering students a more robust understanding of college fit, especially at crucial times during their junior and senior years of high school, students should be able to think in more critical and careful ways about their college selection process.

Second, College Possible sought to engage parents and families more deeply and intentionally throughout the students' junior and senior years. Twice per school year, coaches were to meet with parents and families for sessions to educate and engage them on their students' efforts to find and enroll in a best-fit college. Additionally, College Possible developed and tested methods of ongoing communication with parents and families designed to keep them invested in and apprised of their students' college selection process, while also giving them the opportunity to ask questions of the coaches.

COVID-19 Program Adaptations

The COVID-19 pandemic has affected the entire college process, from admissions to graduation. ACT testing delays, campus visit cancellations, moving admissions deadlines, and other disruptions impacted college admissions and persistence. During the last part of the 2019-20 academic year, College Possible high school programming was adapted to a 100% remote model. College Possible was uniquely equipped to pivot all of their programming to virtual delivery. For example, they have been serving high school students via technology through their Navigate program for six years and have served college students 100% virtually since the program was expanded in 2010 to support college students through college graduation.

College Possible adapted to COVID-19 by adjusting curriculum, dosage, and delivery of the program, as well as referring students to other community providers when needed.

To pivot to 100% virtual programming, College Possible adapted the entire curriculum, as well as dosage and program delivery methods for both juniors and seniors. COVID-19 highlighted the regular and persisting insecurities students face, which were now exacerbated. For example, student engagement now varied even more widely because of less consistent and stable access to technology. In some cases, College Possible offered emergency funds for students to enable technology access. Additionally, College Possible revamped a student resource website so students had access to the information they need to stay on track and meet their individual non-college readiness needs as well as college readiness needs. College Possible also identified alternatives to physical in-person campus visits for their students, since visiting college campuses in person has a profound impact on a student's college decision. Additionally, test preparation programming was disrupted due to college entrance exams, like the ACT/SAT, were being postponed or waived at higher education institutions across the country. This meant only a portion of students were able to complete their entrance exams. As a result, students did not need test preparation support. Due to the demanding needs of the pandemic, program staff and coaches curated and referred students to resources and organizations when their needs fall outside of College Possible's areas of expertise.

Evaluation Design and Methods

Implementation Timeline

The college-fit program was funded for two academic years, the baseline year (2018-19) and year 2 (2019-20). Westat, an independent evaluator engaged by the funder, executed research and/or data-sharing agreements with individual school districts to acquire access to deidentified student-level data. Collection of district administrative data occurred from October to March following the end of each academic year in June. Administrative data was also collected from year prior to implementation, so that student baseline performance prior to their participation in the program could be calculated. Subsequently, impacts were estimated and reported out to external funders.

Evaluation Sample and Participation

College Possible recruited 14 districts to participate in this evaluation: 3 in Pennsylvania and 11 in Minnesota. Smaller districts are more likely to have difficulty in providing administrative data of quality suitable for an impact evaluation. In this case, 3 small districts (of the 14) could not do so and were excluded from the analyses.

On a monthly basis student participation data were submitted from College Possible to Westat at the student-level, indicating how many days student contacts were planned, and how many actually took place. It was possible for students to participate in the program, but for student demographic and academic data to be unavailable due to mobility, and for students to fall out of analyses because they could not be successfully matched or weighted.

In year 1, College Possible served N=2,286 students with a total of 59,158 contact days (an average of 25.9 days per student), and N=845 in Year 2 with a total of 10,144 contact days (an average of 12.0 contact days) as documented by monthly participation reports. A contact day was defined as each day that an individual student had any contact with the program.

Evaluation Research Questions

The evaluation was driven by four research questions:

(1.) What are the impacts on school attendance for students supported by College Possible?
(2.) What are the impacts on high school Grade Point Average (GPA) for students supported by College Possible?
(3.) What are the impacts to on-track to graduation rates (OTI) for students supported by College Possible?
(4.) What change is observed in a measure of college-fit closely aligned to program supports?

The four outcomes of interest are discussed below.

School Attendance (Attendance)

School attendance was used as a generic measure of student engagement in this evaluation. This is because all schools must report it, and it allows for aggregation across districts within a multisite program. While states and districts may offer/require a different number of days, attendance is easily transformed into a percentage of attended (i.e., number of days attended divided by the number of school days offered).

Grade Point Average (GPA)

GPA is the first of two academic outcome measures used in this evaluation. It represents a challenge, as schools and districts measure it in different ways (e.g., a 0–99 scale or an A–F scale (0.0–4.0). To aggregate results across districts, all GPA scales were transformed into a rank order within each district.

On-Track to On-Time Graduation (OTI)

OTI is the second of two academic outcome measures used in the evaluation. This outcome was adopted from the University of Chicago's Consortium of School Research, developed for Chicago Public Schools.[10] The measure was developed specifically for use at the end of 9th grade, to predict whether a student was on track for on-time graduation. As shown in Figure 1, the way the OTI was constructed is directly related to grade progression; therefore, the OTI was used as an outcome for every grade from 9-12. In this way, students from multiple grades could be combined into a single analysis.

The OTI is constructed as follows. First, the number of credits accumulated duringthe academic year is calculated. Second, the number of semester-level failures core course failures is counted. These two variables taken together determine whether a student is considered to be on track or off track for on-time graduation in 4 years (see Figure 1).

Figure 1. On track to graduation indicator matrix

	2+ Core Failures	0-1 Core Failures
<12 Credits	Off-track	Off-track
12+ Credits	Off-track	On-track

College-Fit measure

This is a 19-item self-report measure assessing understanding of different factors that should be considered when judging the personal fit of post-secondary institutions (e.g., *I understand how to determine if a college is a good fit for me*). The measure was collected in both fall and spring semesters and the measure was identical at both collection points. Thirteen items were retained for the analyses presented in this report, all consisting of 5-point scales from *strongly disagree* to *strongly agree*. All items were summed to create a single total score.

Quasi-experimental design to estimate program impacts

The gold standard in evaluation methodology is the randomized-controlled trial (RCT). The use of random assignment of students to a treatment or non-treatment condition can control for unobserved variables, those not captured in district administrative data systems (e.g., home environment). Random assignment allows one to be more confident that observed differences between treatment and control groups are real and potentially replicable.

In this evaluation like most, random assignment was not feasible, therefore Westat used a quasi-experimental (QED) approach to create a peer-comparison group of students statistically similar to College Possible students. Westat used propensity-score weighting (PSW) to ensure that, as a group, College Possible students and their non-participating peers were statistically equivalent at baseline (i.e., before the program began) on demographic variables, attendance, GPA and OTI. This was why prior-year data was collected. Differences at the end of the year could then be attributed to program exposure. However, QED approaches can only control for observed variables – i.e., in this case those available from district data systems.

Impacts were estimated controlling for prior-year academic performance as well as the demographics characteristics provided by school districts.

Findings

Outcome and Impact Analyses

In this section we review findings from four types of analyses. Table 2 summarizes annual findings from analyses utilizing school district data. Table 3 reports primary survey results collected by College Possible.

Positive impacts on OTI rates were found annually regardless of the analytic approach.

The first set of results examines whether program attendance is related to performance at the end of the school year. It is logical that if a program is providing effective supports, more support should be better. We see supporting evidence in the first two rows. Five of the 6 possible outcomes were positive and statistically significant – meaning that students receiving more College Possible contact had higher outcomes at the end of the year

For students in the top 50% of participation, Westat observed positive impacts on all outcomes in the baseline year, but only for OTI during year 2 (COVID-19). A similar, though not identical, finding is observed for students at all levels of participation (i.e., at least one day per year).

Table 2. Annual outcomes and impacts for three analytic strategies

Year	Attendance	GPA	OTI
Relationship between program attendance rate and outcome level			
Baseline Year	↑	⇧	↑
Year 2 – COVID-19	↑	↑	↑
QED impacts for students in the top 50% of participation			
Baseline Year	↑	↑	↑
Year 2 – COVID-19	↓	⇧	↑
QED impacts for students at all levels of participation			
Baseline Year	↑	⇧	↑
Year 2 – COVID-19	↓	⇧	↑

↑, ↓ = Large, bold arrows mean significant and positive or negative relationship.
⇧, ⇩ = Small arrows mean non-significant and positive or negative relationship.
— = No relationship.
N/A = Program did not serve any students in this analytic group.

Table 3 displays the annual outcomes for the student knowledge survey which is closely aligned to College Possible's programmatic activities. This analysis examines *change within program students* from the beginning of the academic year to the end – for knowledge related to college-going and specifically college-fit. In the first year of the evaluation, a positive but non-significant change was observed. In the second year, a positive and statistically significant growth in college-going knowledge was observed. The findings in table 2 indicate that students report increased knowledge and confidence of the components of college-fit in Year 2.

In the following section we report findings from an analysis of cost and benefit-cost based actual inputs and costs collected from College Possible and the impact findings reported in table 2.

Table 3. Annual college-fit survey results

Measure	Year 1 2018-19 Results	Year 2 2019-20 Results	Grades represented
Knowledge of college-fit	⇧	↑	11-12

↑, ↓ = Large, bold arrows mean significant and positive or negative relationship.
⇧, ⇩ = Small arrows mean non-significant and positive or negative relationship.
— = No relationship.
N/A = Program did not serve any students in this analytic group.

Cost analysis and benefit-cost analysis results

The purpose of this portion of the report is to examine the costs and benefits of College Possible programming as part of the evaluation of the College Possible program. Cost analyses were conducted using the ingredients method.[1,2] In the ingredients method, the "costs" of the intervention or program refers to the value of all resources used to generate program impacts, not just those for which money was exchanged. This method considers all costs, regardless of who "pays," whether services/facilities/supplies are provided in kind or through purchase, or how the costs are financed or funded. By using the ingredients method to calculate the total costs of a program, a more thorough and complete picture of the true costs of the program is provided beyond just expenditure and budget data. Likewise, even though the larger evaluation focused on College Possible students' outcomes, the value of those outcomes is spread across all served program students to align with treatment of the program costs.

Additionally, a benefit-cost analysis was conducted for eligible recipients. To be eligible to be included in the benefit-cost analysis, College Possible recipients had to have at least one significant and positive outcome result from the full sample evaluation results.[3] Without a significant, positive outcome result, there is nothing to "monetize." For College Possible both the cost analysis and benefit-cost analysis were conducted.

College Possible provided the ingredients, quantity, unit of measure of the quantity, and unit cost for all ingredients necessary to implement the program as evaluated in the larger College Possible evaluation. This information was provided both in Year 1 and Year 2. Additionally in Year 2 (2019-2020), College Possible also provided qualitative information on the ingredients that had to be changed or stopped because of the COVID-19 pandemic. Westat staff then used the information provided to conduct the cost and benefit-cost analyses reported here.

Results of the cost analyses for Year 1 and Year 2 indicate that while some costs decreased in Year 2 (due to the pandemic), other increased. The overall cost and per student costs for implementation during these two years is displayed in Table 4.

As seen in Table 4, total program costs increased from Year 1 to Year 2. This was due to increases in personnel costs (e.g., personnel time, salary, etc.). Additionally, per student costs also increased due to the reported decrease in students served in Year 2 compared to Year 1.

Looking into the breakdown of costs, the majority of the total program costs in both years of implementation were in the personnel category (e.g., staff salary). A breakdown of the overall costs by cost category is presented in Table 5.

---

[1] Levin, H. M., & McEwan, P. J. (2000). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

[2] Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. Thousand Oaks, CA: Sage Publications.

[3] Full sample evaluation results include attendance, grade point average (GPA), the on-track indicator (OTI), and the aligned outcome.

Table 4. Total program cost and per student costs by year of program implementation

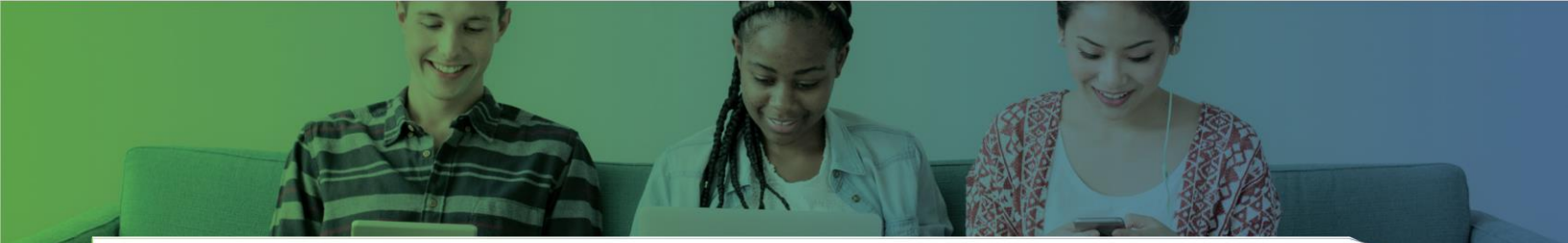
Program year	Total program cost	Total number of students served	Cost per student
Year 1 (2018-19)	$1,507,850	2,286	$660
Year 2 (2019-20)	$1,580,700	845	$1,870

Note. All costs are represented in 2018 dollars and rounded to the nearest $10 to avoid false precision.

Table 5. Breakdown of overall program cost by cost category

Program year	Personnel	Percent personnel	Facilities	Percent facilities	Materials & equipment	Percent materials & equipment	Other	Percent other
Year 1 (2018-19)	$876,790	58	$340,960	23	$264,170	18	$25,930	2
Year 2 (2019-20)	$1,226,350	78	$226,650	14	$116,670	7	$11,030	1

Note. All costs are represented in 2018 dollars and rounded to the nearest $10 to avoid false precision.



The breakdown of costs by cost category shows the increase in personnel costs with smaller decreases in the other cost categories between Year 1 and Year 2. College Possible reported decreases in facilities costs due to the switch to virtual support of, and engagement with, students. Decreases in equipment and materials costs were due to cancellation of events and associated decreases in materials needed. Additionally, decreases in other costs cancellation of other events and a change in reimbursement policies. One sub category within other costs did see a slight increase due to support of connectivity for remote work during the pandemic.

As shown previously, two positive and significant outcomes were found for College Possible in each of the implementation years. In Year 1, the OTI and attendance significant outcomes were monetized. In Year 2, the OTI and aligned outcome[4] were significant. However, only the OTI significant outcome was monetized as the aligned outcome represented the same monetizable outcome of lifetime earnings, which would be a 'double counting' of benefits[5].

In each of the two implementation years included, College Possible demonstrated a positive net benefit. Specifically, both a net present value (NPV) and benefit-cost (BC) ratio were calculated. NPVs larger than 0 can serve as an economic rationale for accepting the investment because it implies the gains are higher than the costs. The NPV of a program can be calculated by subtracting the Present Value (PV) costs from the PV benefits:

$$NPV = B_{PV} - C_{PV}$$

The BC ratio indicates the size of benefits relative to the size of cost. The easiest interpretation of this ratio would be how much benefit is generated by 1 dollar of investment. BC ratios greater than 1 indicate the benefits are higher than the costs. The BC ratio can be obtained by simply dividing the PV benefits by PV costs.[6]

$$BCR = \frac{B_{PV}}{C_{PV}}$$

The results for College Possible across both years indicate a PV and ratio. Results, by year, are presented in Table 6. As seen in Table 6, the NPV is well above 0 and the BC ratio is well above 1. The benefit-cost analysis results suggest the implementation of College Possible included in the present evaluation was a positive investment yielding benefits that are higher than the costs. Specifically in Year 1, the BC ratio suggests that for every $1 spent, a benefit of $14.66 is returned. In Year 2, the BC ratio suggests that for every $1 spent, a benefit of $4.34 is returned. The main driver of the difference in the BC ratios across years is the number of

[4] This is an additional outcome that College Possible selected as part of the evaluation. It was a survey on College Fit.

[5] Decisions were made in calculation of benefits to avoid 'double counting'. For further detail of the monetized values and decisions, please see the Appendix.

[6] Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis.* Thousand Oaks, CA: Sage Publications.

students served. Therefore, the resulting benefits, yield a much higher NPV in Year 1 due to the significantly higher number of students served.

Table 6. Net present value and benefit-cost ratio by year of program implementation

Program year	NPV	BC ratio
Year 1 (2018-19)	$21,303,850	14.66
Year 2 (2019-20)	$5,284,930	4.34

Note. All costs are represented in 2018 dollars and rounded to the nearest $10 to avoid false precision. NPV = Net Present Value; BC = Benefit-Cost

As seen in Table 6, the NPV is well above 0 and the BC ratio is well above 1. The benefit-cost analysis results suggest the implementation of College Possible included in the present evaluation was a positive investment yielding benefits that are higher than the costs. Specifically in Year 1, the BC ratio suggests that for every $1 spent, a benefit of $14.66 is returned. In Year 2, the BC ratio suggests that for every $1 spent, a benefit of $4.34 is returned. The main driver of the difference in the BC ratios across years is the number of students served. Therefore, the resulting benefits, yield a much higher NPV in Year 1 due to the significantly higher number of students served.

Study Limitations

There are two main limitations of the overall evaluation approach. First, the evaluation did not use random assignment. Random assignment of students to College Possible and business-as-usual conditions would allow one to control for unobserved differences between groups. As random assignment was not feasible, a quasi-experimental (QED) design. This means that differences between groups can only be adjusted for statistically on observed variables. The inability to control for unobserved differences in groups means that impact estimates should be interpreted with some caution. For example, if College Possible recruits students on risk factors (e.g., first generation college attendance) that are not represented in the data provided by school districts, it is not possible to take that risk factor into account when creating a peer-comparison group. Second, due to missing data and quality of submitted data on key variables, Westat could not include some program students and comparison students in the analyses conducted to investigate the outcomes of interest. For example, a student missing both the baseline and outcome measure for attendance would be excluded from that analysis. Mobile students are also more likely to be missing data, and more likely to be excluded from the impact estimates for that reason alone. Westat cautions that College Possible students included in program impacts are not necessarily representative of all College Possible students.

Conclusions and Recommendations

This evaluation provided evidence that College Possible's college-fit initiative had positive impacts on student academic outcomes. Most notably, OTI rates were positively impacted in both years of the evaluation, and for both the high intensity and regular intensity student participation groups. This is a substantial achievement given the COVID-19 challenges in the second year. The significant negative impacts on student attendance observed in Table 2 are likely not a serious concern, as tracking of attendance during the beginning of COVID-19 was seriously disrupted and these findings may be due to noise.

As with any QED evaluation, there are reservations due to the fact that students were not randomly assigned to the college-fit intervention. The OTI impacts observed could possibly be observed due to lack of statistical control over unobserved variables A key recommendation for College Possible is that a successful random assignment study could greatly strengthen the evidence for effectiveness of the college-fit initiative.

Appendices

Benefit-Cost Appendix

Monetization of benefits

It is important to note that not every statistically significant result was monetized and included in the benefit-cost analyses. The reason for this is that multiple outcomes can have significant overlap into the same benefit—lifetime earnings. More specifically, when an OTI outcome was significant, a significant GPA outcome was not monetized. Similarly, when an aligned outcome overlapped in content with an OTI outcome (and OTI was significant), the aligned outcome was not monetized. In either case, estimating the benefits of both outcomes would result in "double counting," which would overestimate the potential benefits of the programs.

Specification of benefits and valuation

The following effectiveness evaluation outcomes were identified as "able to be monetized" and were matched to appropriate values, either by finding a competitive market price or equivalent, or by using a shadow pricing technique[7] based on existing research, as described below.

Increased student attendance during a single year of participation was valued through shadow pricing, in other words, by answering the question of how much society is willing to pay for a day of school attendance. For this analysis, state aid for K-12 student attendance was used as a proxy of society's willingness to pay. Note the concept of benefit (or cost) is different from expenditures, and therefore school districts' increased revenue through increased student attendance should not be confused with benefits. Increased student attendance among College Possible students can be valued through how much states value a day of student attendance in the form of state allocations. Note this value will likely be an underestimation of the full value of attendance; at the same time, care must be given to avoid the double-counting of similar benefits from GPA, since attendance may mediate achievement[8]. For this reason, this analysis conservatively accounted for the value of attendance by 50 percent when another achievement outcome was significant.[9, 10] It was further

---

[7] Tower, E., Pursell, G., & Mundial, B. (1986). On shadow pricing. World Bank.

Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. Thousand Oaks, CA: Sage Publications.

[8] Maxwell, L. E. (2016). School building condition, social climate, student attendance and academic achievement: A mediation model. *Journal of Environmental Psychology, 46*, 206-216.

[9] Recommendations can include a range of percentages to adjust estimates to avoid double counting. These usually range from 25–75 percent; therefore, the middle of the range was used here at 50 percent.

[10] Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimar, D. L. (2011). *Cost benefit analysis: Concepts and practice.* New Jersey: Pearson Education.

assumed that (1) the respective states provided approximately 50 percent of per student spending and (2) the average number of days in a school year was 180.

Improved on-track to graduate status during a single year of participation was also valued through shadow pricing. For this analysis, an incremental lifetime benefit value was used[11]. The outcome of additional students on track in relation to the comparison group was also included.

Improved recipient, specific aligned outcomes during a single year of participation was valued through shadow pricing similar to the other outcomes. As College Possible's aligned outcome was the college-fit knowledge survey (including alignment of student preparedness), the most reasonable outcome to value would be continued academic success (i.e., persistence, graduation, etc.). However, this is already taken into account in the outcome monetized for a significant on-track to graduate result. Therefore, although the evaluation found a significant result for this aligned outcome, it was not monetized in addition to the monetized outcome for OTI as this would be a 'double counting' of benefits.

[11] Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. Thousand Oaks, CA: Sage Publications.

Analytic Methods Appendix

A propensity score weighting design was used to create both College Possible and comparison analytic groups. Program estimates were produced using a single-level weighted regression

Propensity score weighting (PSW) design was used to organize the data and produce an estimate of the overall effect on participants' OTI, school attendance, and GPA. This design measures pre-existing differences between groups. However, there will always be reservations about the bias in the estimate of programs' effect on the target outcome because the unmeasured pre-program differences cannot be controlled for and could be confounded (or included) in the estimate. PSW with the average treatment effect on treated (ATT) was used.

After imputing missing data with MI, propensity score, a probability to be assigned to the program group, was calculated for each student. By using the propensity score, a weight was calculated for each student and weighted mean was calculated for baseline measures and demographic variables. If absolute values of the weighted mean difference were less than 0.25 standard deviation, the weighting was considered high quality. When the weighted standardized difference was greater than 0.25 standard deviations for the covariates included in the model, the weighting was considered unacceptable, and the propensity score model was calibrated (i.e., the matching variable set was adjusted and matching rerun) until the acceptable weighted mean differences for the natural pretests were achieved.

This report was authored by Westat, Inc., an independent evaluator. It is abstracted from a more comprehensive evaluation report.

Eric Rolfhus
Emily Diaz
Lauren Decker-Woodrow
Eishi Adachi
Sadeq Sohrabie
Don Barfield
Gay Lamey

Westat, Inc.
1600 Research Blvd,
Rockville, MD 20850
www.westat.com