Working Paper 2022-1
WINTER 2022

Ability Signals and Rigorous Coursework: Evidence from AP Calculus Participation

Christopher Avery, *Harvard University*
Joshua Goodman, *Boston University*

Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

WHEELOCK EDUCATIONAL POLICY CENTER

Ability Signals and Rigorous Coursework: Evidence from AP Calculus Participation

Christopher Avery
Harvard University
Christopher_Avery@hks.harvard.edu

Joshua Goodman
Boston University
edecjosh@bu.edu

February 7, 2022

Abstract

Large racial gaps in rigorous math coursework enrollment are a barrier to improving college completion rates and access to careers in STEM. Perceptions of the math abilities of underrepresented minority students, whether by themselves or their schools, may contribute to such gaps. A regression discontinuity design with data on all Massachusetts public school students shows how receiving a positive ability signal affects subsequent enrollment in advanced high school math coursework. For Black or Hispanic students, particularly those with low reading scores, being labeled as "Advanced" on a 10th grade math test substantially increases the likelihood of later Advanced Placement Calculus course- and exam-taking. White or Asian students and those from higher income high schools show no such effects. Understanding how ability signals affect students' and schools' decisions about assignment to coursework may suggest further ways to address gaps in such enrollment.

1 Introduction

Black and Hispanic students are underrepresented in college STEM majors and in subsequent jobs in STEM fields (Fry et al., 2021), including in medical school (Morris et al., 2021). Patterns of enrollment in rigorous STEM coursework in high school may explain some of these gaps. Descriptive evidence suggests high school students who take Advanced Placement (AP) STEM courses are substantially more likely to complete a STEM college degree than observationally similar peers who do not take such courses (Smith et al., 2018). Recent causal evidence suggests that taking AP coursework can boost students' interest in pursuing a STEM major in college (Conger et al., 2021) and that earning AP credits can increase the number of college STEM courses completed (Gurantz, 2021).

There exist, however, clear disparities by race and income in enrollment in such advanced coursework. The most recent data available suggests that Black students comprise 15% of each year's graduating high school class but only 9% of AP exam takers and less than 5% of those who earn a passing score on at least one AP exam.[1] Such gaps are not as stark for Hispanic and Latino students, who are 19% of the graduating class but only 17% of those who earn at least one passing score. Gaps by income are arguably even larger. Whereas 48% of K-12 students in the United States are eligible for free or reduced-price lunch, but those students represent only 27% of AP exam takers and only 22% of those with at least one passing score. Such gaps are not limited to Advanced Placement coursework. A 2014 report from the Civil Rights Division of the US Department of Education shows that Black and Hispanic students are similarly underrepresented in elementary school gifted and talented programs, as well as in core and advanced quantitative courses in middle and high school.[2]

In this paper, we study one potential determinant of enrollment in advanced STEM coursework in high school, namely the signal sent by the label used to describe a student's performance

[1] "10th Annual AP Report to the Nation," College Board, 2014. Available at http://media.collegeboard.com/digitalServices/pdf/ap/rtn/10th-annual/10th-annual-ap-report-to-the-nation-single-page.pdf.

[2] "Data Snapshot: College and Career Readiness Report," U.S. Department of Education, 2014. Available at https://www2.ed.gov/about/offices/list/ocr/docs/crdc-college-and-career-readiness-snapshot.pdf.

on an important standardized exam. Students scoring in the top category of Massachusetts' standardized 10th grade math test are described as "Advanced", a positive signal that roughly half of students receive in any given year. We use a regression discontinuity design to compare future course-taking and performance for students just above and below the threshold that distinguishes that category from the "Proficient" one below it. Students just above and below this threshold are similar in terms of demographics and other characteristics, differing only in whether they receive this positive signal. This allows us to identify the causal impact of that signal on subsequent course-taking outcomes.

We first document large gaps in AP Calculus course-taking by race/ethnicity and by income. Black or Hispanic students are 8 percentage points less likely to take AP Calculus than their White peers, a gap largely unchanged when comparing students in the same high school. Low income students are 9 percentage points less likely to take AP Calculus than their non-low income peers, or 5 percentage points less than those in their same high school. We then show that, for Black or Hispanic students, being labeled as "Advanced" on the 10th grade math test substantially increases the likelihood of later Advanced Placement Calculus course- and exam-taking. The effect is larger among students with low reading scores, suggesting that the signal carries greater weight for students without other markers of high achievement. The "Advanced" threshold is near the median of the math ability distribution, so that the marginal student induced to enroll tends to earn low grades in the AP Calculus course and low scores on the AP Calculus exam. The enrollment of White or Asian students is unaffected by this ability signal. The effect of these signals on low income students generally is about half the magnitude as the effect on Black and Hispanic students and only marginally statistically significant. We see little evidence of heterogeneity by gender, perhaps related to the lack of a gender gap in AP Calculus participation rates.

These findings contribute to two existing strands of research. First, and most narrowly, a growing body of literature documents that preliminary signals of performance can have substantial effects in directing the focus of future course choices by students. Papay et al. (2016) find that low-income urban students who score just above a scaled score performance threshold on a standardized state high school test are significantly more likely to enroll in college than students who score just below the threshold. Avery et al. (2018) find that students with raw scores just over the threshold for AP college credit are significantly more likely to choose a college major corresponding to the subject of that test than are students with raw scores just below that same threshold. Goldin (2015) finds that women respond more negatively than men to less than perfect grades in an introductory economics course. Rogers and Feller (2016) find that students randomly selected to review the work of academically stronger classmates are less likely to complete their large scale on-line course, perhaps discouraged by their lowered perception of their own skills. Our paper adds to the evidence that positive signals can change students' coursework decisions, including as early as high school.

Second, and more broadly, recent studies find clear evidence of bias in the selection of students for advanced coursework. Card and Giuliano (2016) observe that the introduction of a universal screening test for admission to a gifted education program significantly increased the proportion of minority and low-income students in the program. Similarly, Dougherty et al. (2015) find that the representation of minority and low-income students in Algebra I in 8th grade increased when the Wake County School District introduced a formal rule to select students for the course on the basis of prior achievement. From test scores and grades to less formal feedback, students receive a stream of noisy signals about their academic performance and perceived ability levels throughout their years in K-12 classrooms. The results of these prior studies suggest that traditionally disadvantaged students receive, on balance, more negative signals than others conditional on past test scores and grades. Our results suggest that positive signals can thus affect the decisions of such students, or perhaps their teachers.

We conclude by discussing potential mechanisms and implications of our findings. We note that our observed effects may be driven by changing the course choices of students and their families, or of teachers and school counselors who assign students to courses (Francis et al., 2019; Mulhern, 2020). We argue that, though this particular ability signal was not well-targeted toward students likely to succeed in AP Calculus courses, states and school districts could generate more informative signals designed to identify students appropriate for a given advanced course. Such signals could improve participation in appropriately rigorous coursework.

2 Data and Empirical Strategy

We use administrative data from the Massachusetts Department of Elementary and Secondary Education for public school students whose projected high school graduation years are 2011-14. For each student, we observe four important pieces of data: demographic information; 10th grade standardized test scores; high school transcripts; and AP exam results. The demographic information includes students' gender, race/ethnicity, and low income status (measured by receipt of federally subsidized school meals).

2.1 Outcomes

The high school transcripts and AP exam results provide our main outcomes. For each student, the transcripts provide a record of each high school course taken and grade earned in that course. We observe both course titles and course codes standardized by the state, making it easy to identify AP courses. Our main course outcome of interest is an indicator that a student took any AP Calculus class, including either or both of the College Board's two AP Calculus classes, AB and BC Calculus. According to the College Board, AB Calculus is equivalent to a semester of calculus at most colleges, while BC Calculus includes additional topics and is equivalent to a year of college calculus. We measure students' success in those courses by defining an indicator for earning at least a B in any AP Calculus course.

The AP exam results, linked to individual Massachusetts students by the College Board, provide an additional measure of participation and success with college-level math. Not all students who enroll in AP Calculus courses end up taking the corresponding AP exam, so we define an indicator for taking any AP Calculus exam. We also define an indicator for "passing" the AP Calculus exam by receiving at least a 3 on the College Board's 1-5 grading scale. Colleges typically grant credit only for AP scores of 3 or higher, which we therefore interpret as having learned enough to meet college-level expectations.

2.2 10th grade test scores

We then ask how students' 10th grade standardized math scores affect their enrollment and performance in AP Calculus courses and exams. All Massachusetts public high school students take the Massachusetts Comprehensive Assessment System (MCAS) in 10th grade. The MCAS has been subsequently updated so we describe here the "Legacy" MCAS taken by the cohorts we study. This MCAS contained three sections: Mathematics, English Language Arts (ELA), and Science and Technology/Engineering (STE). Students took the exam in spring of 10th grade and results were distributed to schools and families in the fall of 11th grade, roughly a month or two after the school year had started. MCAS results in each subject were reported as scaled scores ranging from 200 to 280 in increments of 2 points. Importantly, these scaled scores were also converted into a "Performance Level", which could fall into four possible categories. Students were "Advanced" if they scored 260-80, "Proficient" if they scored 240-258, "Needs Improvement" if they scored 220-238, and "Failing" if they scored 200-218.

The Performance Level assigned to each student may have been particularly salient to students and schools for three reasons. First, students in these cohorts needed to score at least Needs Improvement on both subjects to graduate from high school. The threshold between Needs Improvement and Failing thus involved very high stakes for students and schools. Second, eligibility for merit-based aid called the Adams Scholarship is determined in part by whether a student scored Advanced on one subject and at least Proficient on the other. That aid was worth up to about $7,000 in total over four years and did affect some students' decisions about where to enroll in college (Goodman, 2008; Cohodes and Goodman, 2014). The Adams Scholarship was the only formal stakes associated with the threshold between Advanced and Proficient. Third, after a cover page, the "Parent/Guardian Report" mailed to families in the fall following the 10th grade MCAS listed various aspects of the student's MCAS performance, with the performance level featured prominently at the top of the page above the scaled score itself.[3] The Performance Level is thus the first evidence students and families see when they read the test report and the one that directly contributes to a student's high school graduation and merit aid eligibility.

[3] See Figure A.1 for the template of the report mailed to these cohorts.

2.3 Summary statistics and course-taking gaps

Before turning to estimation of ability signal effects, we describe the data. As Table 1 shows, among the overall high school classes of 2011-14 whom we study, 11.3 percent took at least one AP Calculus course and 7.3 percent earn at least a B in such a course. Similarly, 10.8 percent of students took at least one AP Calculus exam and 7.6 percent earned at least a 3 on such an exam. It thus seems that the vast majority of students who take an AP Calculus course also take the corresponding exam. Only two-thirds of such students appear, however, to succeed in such courses and exams.

Dividing the sample into those who did and did not score "Advanced" on the 10th grade math MCAS reveals a few important facts. First, nearly half of students during this time period achieved an "Advanced" score, suggesting that the threshold falls roughly at the median of the math ability distribution. Since less than one-eighth of students take an AP Calculus exam, a typical student on the margin between "Advanced" and "Proficient" is an unlikely candidate for AP Calculus. Second, and closely related, AP Calculus course and exam taking is almost entirely concentrated among the "Advanced" students. Over 20 percent of such students take AP Calculus courses and exams, while fewer than one percent of students below the "Advanced" threshold do so. Third, compared to the student population of Massachusetts, "Advanced" students are less likely to be Black, Hispanic, or low income.

In Table 2, we measure the extent of and statistical explanations for demographic disparities in AP Calculus course-taking rates. The first three columns study raw gaps in such rates. Compared to their largely White or Asian counterparts, Black or Hispanic students are 8.4 percentage points less likely to take AP Calculus courses. Low income students are 8.6 percentage points less likely to take AP Calculus courses than their higher income peers. Gender gaps in course-taking are very small, with female students only 0.4 percentage points less likely to take AP Calculus courses than their male peers. Column 4, which includes all of these covariates simultaneously, suggests that conditioning on student income reduces the Black/Hispanic gap in course-taking rates to 5.0 percentage points.

These differences in AP Calculus course-taking appear driven more by differences in math ability as of 10th grade than by high school-level differences. Adding high school fixed effects to the regression model, as in column 5, has little effect on demographic differences in course-taking rates. Black or Hispanic students are 7.1 percentage points less likely to take AP Calculus than their largely White or Asian counterparts within the same high school. Across high school variation in course offerings therefore seems unlikely to explain such differences. Controlling for math ability, as measured by 10th grade MCAS scores, statistically explains more of these gaps. As seen in column 6, when comparing students from the same high school and of the same math ability, the gap in AP Calculus course-taking rates shrinks to 1.3 percentage points by race and 0.8 percentage points by income, and becomes small but positive for women compared to men. These raw demographic disparities in course-taking rates and the disparities in underlying test scores that partly explain them both motivate our regression discontinuity approach.

2.4 Regression discontinuity

We hypothesize that the Performance Level reported for each student constitutes a potentially important signal of that student's mathematical knowledge or ability, one that is more salient or easily understood than the scaled score itself. That signal may affect the decision-making of the students, their families, or their teachers and guidance counselors who are partly responsible for assigning students to course schedules. Simple comparisons of the coursework patterns of students with different performance levels would not yield causal estimates of those signals because of large differences in the underlying abilities of those students.

For this reason, we use a regression discontinuity approach to test whether these ability signals affect students' participation in AP Calculus. Specifically, we compare AP course enrollment and course grades, as well as AP exam taking and scores, for students whose 10th grade MCAS math scores place them just above and below the threshold between "Advanced" and "Proficient". Such students should be nearly identical in terms of underlying mathematical ability and other baseline characteristics and should differ only in the Performance Level they have been labeled with. Any differences in AP Calculus participation should therefore be attributable to the signal itself.

The logistics of the MCAS test seem to fit the requirements for the validity of a regression discontinuity approach. The scoring rubric that translates raw scores into scaled scores changes from year to year, so that the raw score version of the thresholds are not known at the time of the test. The exam includes many questions, so it is unlikely that any student would be able to target a particular raw score with any degree of precision. The exam is also centrally graded, so there appears to be little scope for deliberate manipulation of scores.

We therefore run regression discontinuity models of the following form:

$$APCalc_{isc} = \beta_0 + \beta_1 Advanced_{isc} + \beta_2 D_{isc} + \beta_3 Advanced * D_{isc} + \delta_{sc} + \epsilon_{isc}$$

Here, the outcomes are various indicators of participation in AP Calculus for student i from high school s and graduating in cohort c. This model includes an indicator for achieving the Advanced Performance Level ($Advanced$), the distance of a student's raw math MCAS score from the Advanced threshold (D), and the interaction of those two. We default to a bandwidth of 4 raw score points so that this local linear approach essentially fits straight lines over a narrow window on either side of the Advanced threshold, so that the coefficient of interest (β_1) estimates the difference in the intercepts of those lines. Inclusion of high school by graduating cohort fixed effects δ implies that the coefficient is interpretable as the difference in AP Calculus participation between two high school classmates of the same mathematical ability who differ only in the ability signal they received.

If the regression discontinuity assumptions hold, β_1 represents the causal effect of the "Advanced" signal on students' participation in AP Calculus. We provide two pieces of evidence that these regression discontinuity assumptions are likely satisfied. First, the density of the running variable is quite smooth across the entire sample and across sub-groups we examine.[4] This is consistent with students having little scope for manipulation of the running variable. Second, baseline covariates such as gender, race/ethnicity, and low income status all vary smoothly across the threshold. This implies that the students being compared by the regression discontinuity model do not differ at baseline in any way beyond the ability signal they receive.

We note that it is not obvious that all students, families or teachers will attach importance

[4]See Figure A.2 for details.

to Performance Levels. Since there is a well-known and simple rule to convert scaled scores to Performance Levels, some may correctly interpret the difference between a scaled score of 258 and of 260 to be of minimal interest, despite that these scores are associated with different Performance Levels. If so, this response would only serve to minimize the difference in behavior of students just above and below this cutoff. For this reason, we view our empirical analysis as a conservative test of the effect of an unexpected positive signal on future course choice and performance.

3 Results

3.1 AP Calculus Engagement

The Advanced signal has little effect on AP Calculus course-taking for the average Massachusetts high school student but meaningfully increases enrollment rates for Black or Hispanic students. Panels A and B of Figure 1 shows this graphically. We see a clear discontinuity in AP Calculus course-taking for Black and Hispanic students, but no such discontinuity for the overall sample. The panels also show that only three to four percent of students around the Advanced threshold enroll in AP Calculus courses, a pattern that holds both overall and for the Black or Hispanic subsample. The Advanced threshold is low enough that students near it do not typically take AP Calculus courses. Interestingly, at this point of the ability distribution, Black or Hispanic students are somewhat more likely than the average student to be enrolled in AP Calculus for any given 10th grade math score.

Regression-based estimates of the visually observed discontinuity confirm that the Advanced signal substantially increases Black or Hispanic students' AP Calculus course-taking rates. As panel A of Table 3 shows, the Advanced signal has no effect on AP Calculus course-taking for the average Massachusetts student or for White or Asian students specifically. The Advance signal does, however, raise enrollment rates of Black or Hispanic students by 2.5 percentage points. This represents a doubling of AP Calculus course-taking rates relative to the low baseline rate for those just below the threshold.

Similar to AP Calculus course-taking, the Advanced signal has little effect on AP Calculus exam-taking for the average student but meaningfully increases exam-taking rates for Black or Hispanic students. Panels C and D of Figure 1 shows this graphically. We see a discontinuity in AP Calculus exam-taking for Black or Hispanic students but no such discontinuity for the overall sample. The panels also show that few students around the Advanced threshold take AP Calculus exams. The Advanced threshold is low enough in the ability distribution that it is quite rare for students near it to take AP Calculus exams. Again, at this point of the ability distribution, Black or Hispanic students are somewhat more likely than the average student to take an AP Calculus exam for any given 10th grade math score.

Regression estimates confirm the visual evidence of the Advance signal's impact on AP exam-taking for Black or Hispanic students, of a very similar magnitude to the course-taking impacts. As the second row in panel A of Table 3 shows, the Advanced signal has little impact on the AP exam-taking rates of the White or Asian students but raises those of Black or Hispanic students by 2.4 percentage points. This again represents a doubling relative to the baseline exam-taking rate below the threshold of less than two percentage points. Both course-taking and exam-taking measures suggest Black or Hispanic students' engagement with AP Calculus is meaningfully increased by a positive ability signal.

3.2 Heterogeneity and AP Calculus Success

Among Black or Hispanic students, the Advanced signal has the greatest effect on AP Calculus engagement for those less likely to have received other positive ability signals. Columns 4 and 5 of Table 3 divide Black or Hispanic students into those whose 10th grade ELA scores are below or above the state average for that cohort. The effects of a positive math ability signal are 2-3 times larger for Black or Hispanic students with lower reading scores, for both course- and exam-taking. A positive ability signal may be more salient to the students, their families or their teachers, when the student has not received other such positive signals.

The marginal student induced to engage with AP Calculus by course or exam does not appear to succeed at either. As panel B of Table 3 shows, the Advanced signal has no impact on the proportion of students earning a B or better in AP Calculus. That this holds true for Black or

Hispanic students implies that nearly all of the marginal students induced to enroll earned a grade of C or worse. Consistent with the already low AP Calculus enrollment rates near the threshold, this suggests that students on the margin between Proficient and Advanced (at the median of the distribution given the thresholds' location for these cohorts) are not sufficiently well-prepared to succeed in such coursework. Consistent with coursework results, the marginal student induced to take the AP Calculus exam does not perform well on it. The Advanced signal has no impact on the proportion of students earning a 3 or higher on the exam.[5] This implies that all of the marginal students induced to take the exam earned a score of 1 or 2, too low to earn college credit. Students on the margin between Proficient and Advanced do not appear sufficiently well-prepared to succeed on AP Calculus exams. This finding is consistent with a recent experimental study of the introduction of AP science classes in 23 high schools, where few students were well prepared for such courses and only a small proportion offered an AP science course were able to do well on the exam (Conger et al., 2020).

Heterogeneity by race/ethnicity in the impact of these ability signals on AP Calculus engagement is substantially larger than heterogeneity by other student characteristics. Table 4 splits the sample by student income and gender. The Advanced signal has no clear impact on non-low income students and an impact on low income students that is only marginally significant and half the magnitude of the impact on Black or Hispanic students. Heterogeneity by race/ethnicity is thus stronger than by income. Ability signals have no clear impact on AP Calculus engagement for the average male or female student and even among Black or Hispanic students estimated impacts are similar and statistically indistinguishable by gender. This is arguably consistent with our earlier evidence of nearly no gender gap in AP Calculus engagement.

Though we are primarily focused on AP Calculus coursework, which is designed to provide college credit when students pass the associated AP exam, many students in our data appear to take Calculus courses that do not follow the AP curriculum. The positive ability signal appears to have no effect on the taking of such courses, both overall and for Black or Hispanic students

[5] The statistically significant negative coefficient for the overall sample and White or Asian subsample is both very small and sensitive to regression specification.

specifically.[6] These courses are 3-4 times more common than AP Calculus courses for students near the Advanced threshold, implying that they require less academic preparation than the AP version. The ability signal appears to matter more for highly selective coursework than that open to wider range of students.

3.3 Robustness and Placebo Tests

The estimated impacts of ability signals on AP Calculus engagement is robust to changing our baseline regression discontinuity specification. We do so by: adding demographic controls for gender, race and low income status; doubling the bandwidth and fitting quadratics on either side of the threshold; and fitting quadratics to the complete data set, a so-called global RD.[7] Across all of these alternate specifications, effects on Black or Hispanic students' course-taking always remain at statistically significant and between 1.9 and 2.7 percentage points in magnitude. Effects on exam-taking also remain significant and between 2.1 and 2.4 percentage points. Tests of other nearby "placebo" thresholds show that estimated treated effects are largest and most statistically significant at the true threshold, suggesting our results are not a statistical artifact.[8]

Tests of placebo outcomes and treatments also strengthen our argument for a causal connection between math ability signals and AP Calculus participation. An Advanced signal from the ELA portion of the 10th grade MCAS has no effect on AP Calculus course- or exam-taking, even when the sample is limited to Black or Hispanic students.[9] The Advanced signal from the math portion of the MCAS does not increase the number of AP courses taken other than AP Calculus.[10] Though achieving the advanced performance level in one subject test could inspire a student to be more academically ambitious in other subjects, it seems natural to expect a larger effect of a positive result on the MCAS test within subject than across subjects. Consistent with the simple model that test scores and their corresponding labels convey domain-specific information about academic strength, math ability signals appear to affect only math coursework and non-math ability signals

[6]See Table A.1 for details.
[7]See Table A.2 for details.
[8]See Figure A.3 for details.
[9]See panel A of Figure A.3 for details.
[10]See panel A of Figure A.3 for details.

have no effect on math coursework.

4 Discussion

Though a positive ability signal in math clearly increases Black or Hispanic students' participation in AP Calculus, our data does not help us identify the precise channel through which this effect happens. We see two main possibilities, that such signals change decisions made by students and their families or that they change decisions made by schools and teachers (or both). The signal may change students' or families' perceptions of students' abilities, motivating them to proactively seek out enrollment in AP Calculus coursework. Alternatively, even without action by students or families, the teachers or school counselors in charge of course assignment may use MCAS performance levels as a heuristic when making such decisions. Recent work suggests that school counselors can have meaningful impacts on the probability of AP test-taking (Mulhern, 2020) and can exhibit racial and gender biases when assigning otherwise identical students to such coursework (Francis et al., 2019). Conditional on other indicators of preparation, the "Advanced" label may give a boost to educators' perceptions of a student's readiness for rigorous coursework. This label may be particularly salient for assignment to "Advanced Placement" coursework given that such course titles contain the same word as the signal itself.

That the ability signal effects we document are strongest for Black or Hispanic students suggests they or their educators require additional evidence before considering those students academically prepared for rigorous coursework. Consistent with larger effects on students who did not receive an Advanced signal in English Language Arts, Black or Hispanic students may be less accustomed to receiving affirmation of their skills. Their teachers and guidance counselors may also, due to implicit or explicit bias, treat such signals as new information, even when a student's wider academic record likely implies the Advanced label carries little such information. That the signal affects Black and Hispanic students more strongly than low income students could be driven by teachers and counselors looking for way to increase racial diversity in AP Calculus participation. Alternatively, race and ethnicity may be more strongly correlated than subsidized meal status with parental education or other aspects of family life that affect such coursework choices.

Finally, our results suggest states and school districts should consider devising signals to students and educators that are clearer and more predictive of success in the targeted outcome. The particular margin studied here was not well targeted toward AP Calculus preparation, as it was set low enough in the math distribution that the marginal student did not flourish in AP Calculus. States and school districts could, however, use test scores and past coursework and grades to devise more accurate measures of suitability for advanced coursework. They could then label students more appropriately with categories related to specific courses. Our results suggest that better targeted heuristics have the potential to improve the number and diversity of students enrolling in rigorous coursework.

References

Avery, C., O. Gurantz, M. Hurwitz, and J. Smith (2018). Shifting college majors in response to advanced placement exam scores. *Journal of Human Resources 53*(4), 918–956.

Card, D. and L. Giuliano (2016). Universal screening increases the representation of low-income and minority students in gifted education. *Proceedings of the National Academy of Sciences 113*(48), 13678–13683.

Cohodes, S. R. and J. S. Goodman (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics 6*(4), 251–85.

Conger, D., A. I. Kennedy, M. C. Long, and R. McGhee (2021). The effect of advanced placement science on students' skills, confidence, and stress. *Journal of Human Resources 56*(1), 93–124.

Conger, D., M. C. Long, and R. McGhee Jr (2020). Advanced placement and initial college enrollment: Evidence from an experiment. *Annenberg Working Paper Series 20*(340).

Dougherty, S. M., J. S. Goodman, D. V. Hill, E. G. Litke, and L. C. Page (2015). Middle school math acceleration and equitable access to eighth-grade algebra: Evidence from the wake county public school system. *Educational Evaluation and Policy Analysis 37*(1_suppl), 80S–101S.

Francis, D. V., A. C. de Oliveira, and C. Dimmitt (2019). Do school counselors exhibit bias in recommending students for advanced coursework? *The BE Journal of Economic Analysis & Policy 19*(4).

Fry, R., B. Kennedy, and C. Funk (2021). Stem jobs see uneven progress in increasing gender, racial and ethnic diversity. *Pew Research Center Report, April 1*.

Goldin, C. (2015). Gender and the undergraduate economics major: Notes on the undergraduate economics major at a highly selective liberal arts college. *manuscript, April 12*.

Goodman, J. (2008). Who merits financial aid?: Massachusetts' adams scholarship. *Journal of public Economics 92*(10-11), 2121–2131.

Gurantz, O. (2021). How college credit in high school impacts postsecondary course-taking: The role of advanced placement exams. *Education Finance and Policy 16*(2), 233–255.

Morris, D., P. Gruppuso, H. A. McGee, A. L. Murrillo, A. Grover, and E. Y. Adashi (2021). Diversity of the national medical student body — four decades of inequities. *New England Journal of Medicine 384*, 1661–1668.

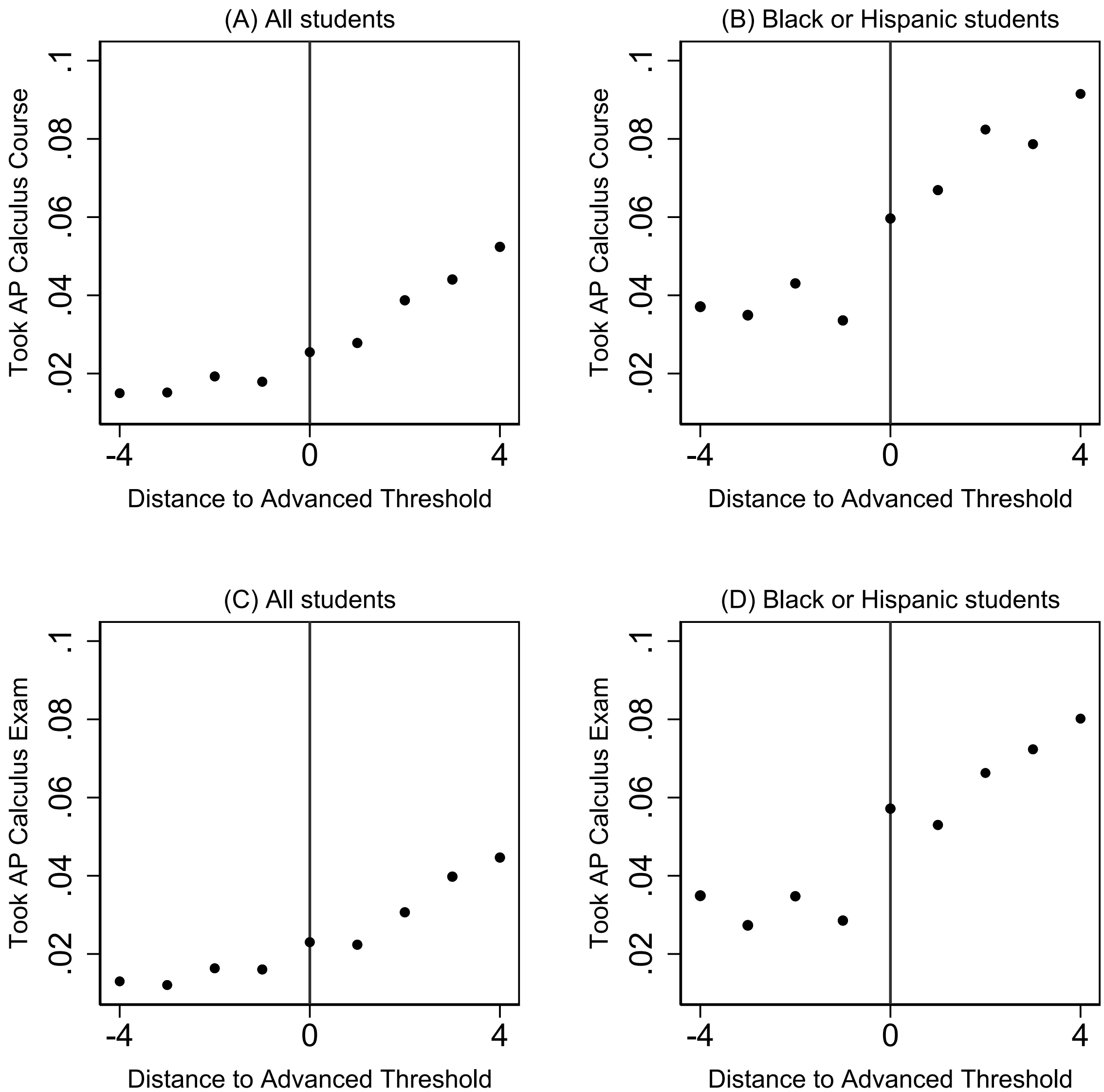
Mulhern, C. (2020). Beyond teachers: Estimating individual guidance counselors' effects on educational attainment.

Papay, J. P., R. J. Murnane, and J. B. Willett (2016). The impact of test score labels on human-capital investment decisions. *Journal of Human Resources 51*(2), 357–388.

Rogers, T. and A. Feller (2016). Discouraged by peer excellence: Exposure to exemplary peer performance causes quitting. *Psychological science 27*(3), 365–374.

Smith, K., S. Jagesic, J. Wyatt, and M. Ewing (2018). Ap® stem participation and postsecondary stem outcomes: Focus on underrepresented minority, first-generation, and female students. *College Board*.

Figure 1: AP Calculus Participation by Distance from "Advanced"



Notes: Shown above are mean proportion of Massachusetts high school students who took any Advanced Placement Calculus course (panels A and B) or exam (panels C and D), by the distance of their 10th grade math MCAS score from the threshold defining the "Advanced" category.

Table 1: Summary Statistics

	All	10th grade math signal: Advanced	10th grade math signal: Not Advanced
	(1)	(2)	(3)
(A) Demographics			
Male	0.506	0.508	0.504
White	0.708	0.792	0.624
Black	0.087	0.047	0.127
Hispanic	0.131	0.066	0.195
Asian	0.052	0.074	0.030
Low income	0.311	0.185	0.437
(B) Course/exam-taking			
Took AP Calculus course	0.113	0.220	0.006
Earned A or B in AP Calculus	0.073	0.144	0.001
Took AP Calculus exam	0.108	0.211	0.005
Passed AP Calculus exam	0.076	0.152	0.000
N	284,255	141,925	142,330

Notes: Mean values of key variables are shown for the Massachusetts public high school classes of 2011-14. Column 1 includes all students. Columns 2 and 3 divide the sample into those who did and did not score "Advanced" on the 10th grade math portion of the MCAS. In panel B, AP Calculus includes both AB Calculus and BC Calculus.

Table 2: AP Calculus Course-Taking Gaps

	(1)	(2)	(3)	(4)	(5)	(6)
Black/Hispanic	-0.084***			-0.050***	-0.071***	-0.013***
	(0.006)			(0.005)	(0.006)	(0.003)
Low income		-0.086***		-0.065***	-0.048***	-0.008***
		(0.006)		(0.006)	(0.004)	(0.002)
Female			-0.004**	-0.004*	-0.008***	0.006***
			(0.002)	(0.002)	(0.002)	(0.001)
Constant	0.133***	0.141***	0.117***	0.147***	0.149***	-0.012***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.002)	(0.003)
N	279,615	279,615	279,615	279,615	279,615	279,615
High school fixed effects					Y	Y
Math score controls						Y

Notes: Heteroskedasticity robust standard errors clustered by high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each estimate comes from regressing an indicator for taking an AP Calculus course on the variables shown. The sample includes all students whose 10th grade MCAS timing implies they are in the high school classes of 2011 through 2014. Column 5 also includes high school fixed effects, while column 6 controls for a cubic function of 10th grade MCAS math score.

Table 3: AP Calculus Engagement and Success

	All students (1)	White/ Asian (2)	Black/ Hispanic (3)	Black/Hispanic: Low ELA scores (4)	Black/Hispanic: High ELA scores (5)
(A) AP Calculus Engagement					
Took AP Calculus course	0.002	-0.001	0.025***	0.043***	0.015
	(0.003)	(0.003)	(0.009)	(0.012)	(0.014)
Control mean	0.018	0.022	0.021	0.024	0.023
Took AP Calculus exam	0.001	-0.003	0.024***	0.033***	0.013
	(0.003)	(0.002)	(0.008)	(0.011)	(0.012)
Control mean	0.016	0.021	0.018	0.021	0.020
N	63,360	51,343	10,427	4,303	5,711
(B) AP Calculus Success					
A or B in AP Calculus	-0.001	-0.001	0.003	0.004	0.000
	(0.001)	(0.001)	(0.004)	(0.005)	(0.007)
Control mean	0.006	0.007	0.006	0.007	0.008
Passed AP Calculus exam	-0.002**	-0.002***	0.002	0.001	-0.001
	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)
Control mean	0.002	0.002	0.002	0.002	0.002
N	63,360	51,343	10,427	4,303	5,711

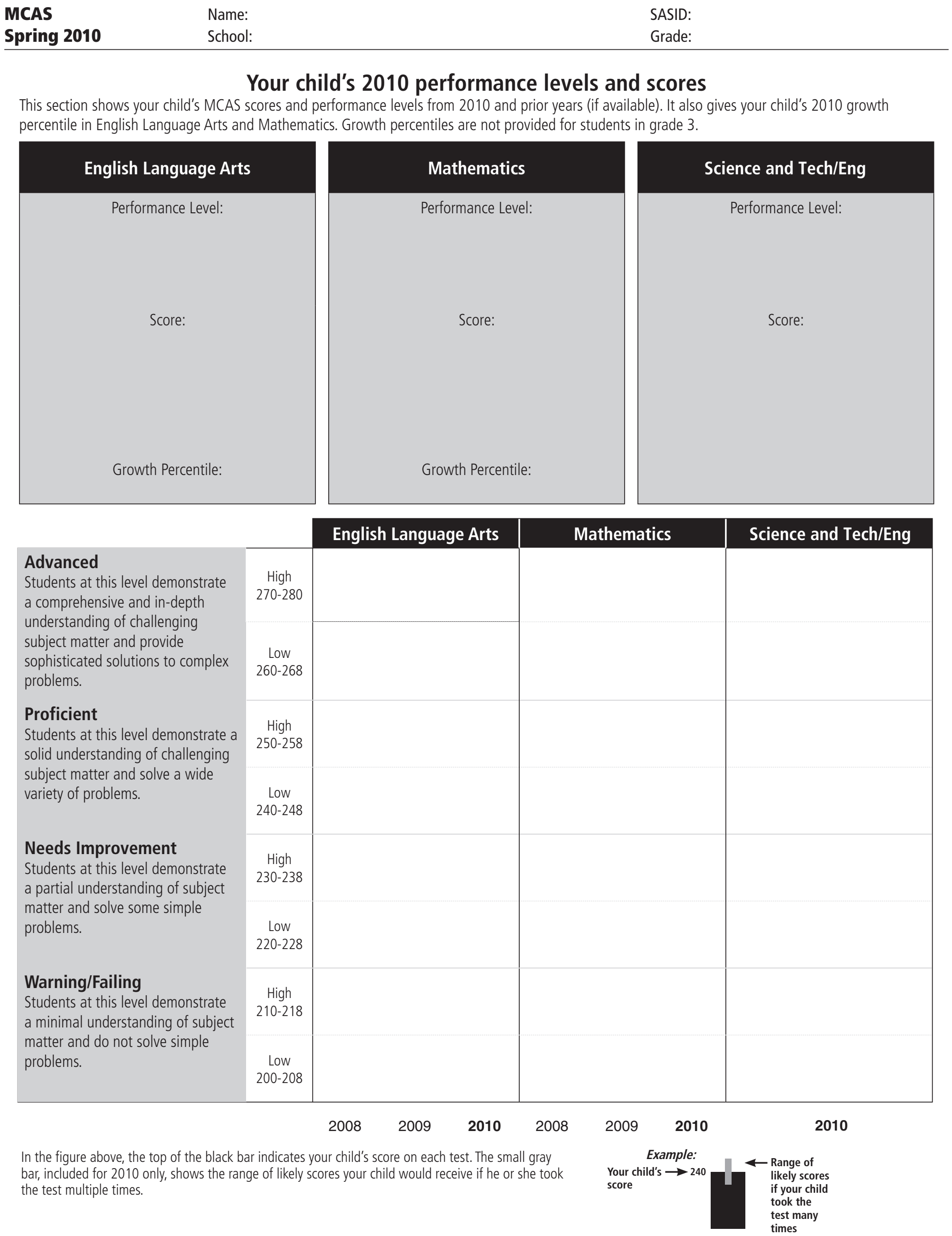
Notes: Heteroskedasticity robust standard errors clustered by high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each estimate comes from a local linear regression discontinuity model that regresses an indicator for a given outcome on an indicator for receiving an score on the 10th grade math MCAS, raw score distance from that threshold, and the interaction of the two. All regressions include high school by graduating class fixed effects and use a bandwidth of four raw score points. The sample includes all students whose 10th grade MCAS timing implies they are in the high school classes of 2011 through 2014. Outcomes are defined as taking an AP Calculus course or exam (panel A) and earning at least a B in an AP Calculus exam or passing an AP Calculus exam (panel B) by the end of 12th grade. Columns 4 and 5 limit the sample to Black or Hispanic students with ELA scores below or above the statewide average for that cohort. Listed below each column is the mean of the outcome for those one raw score point below the threshold.

Table 4: Heterogeneity by Income and Gender

	Non-low income	Low income	Male	Female	Black/Hispanic Male	Black/Hispanic Female
	(1)	(2)	(3)	(4)	(5)	(6)
Took AP Calculus course	0.000	0.012*	0.002	0.002	0.022*	0.031**
	(0.003)	(0.007)	(0.003)	(0.004)	(0.012)	(0.014)
Control mean	0.021	0.022	0.021	0.022	0.024	0.022
Took AP Calculus exam	-0.001	0.010	0.001	0.001	0.023*	0.026**
	(0.003)	(0.006)	(0.003)	(0.003)	(0.012)	(0.013)
Control mean	0.019	0.020	0.019	0.020	0.021	0.020
N	46,687	16,517	30,851	32,419	4,907	5,182

Notes: Heteroskedasticity robust standard errors clustered by high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each estimate comes from a local linear regression discontinuity model that regresses an indicator for a given outcome on an indicator for receiving an score on the 10th grade math MCAS, raw score distance from that threshold, and the interaction of the two. All regressions include high school by graduating class fixed effects and use a bandwidth of four raw score points. The sample includes all students whose 10th grade MCAS timing implies they are in the high school classes of 2011 through 2014. Outcomes are defined as taking an AP Calculus course or exam by the end of 12th grade. Listed below each column is the mean of the outcome for those one raw score point below the threshold.

Figure A.1: Sample Family/Guardian MCAS Report



MCAS
Spring 2010

Name:
School:

SASID:
Grade:

Your child's 2010 performance levels and scores

This section shows your child's MCAS scores and performance levels from 2010 and prior years (if available). It also gives your child's 2010 growth percentile in English Language Arts and Mathematics. Growth percentiles are not provided for students in grade 3.

English Language Arts	Mathematics	Science and Tech/Eng
Performance Level:	Performance Level:	Performance Level:
Score:	Score:	Score:
Growth Percentile:	Growth Percentile:	

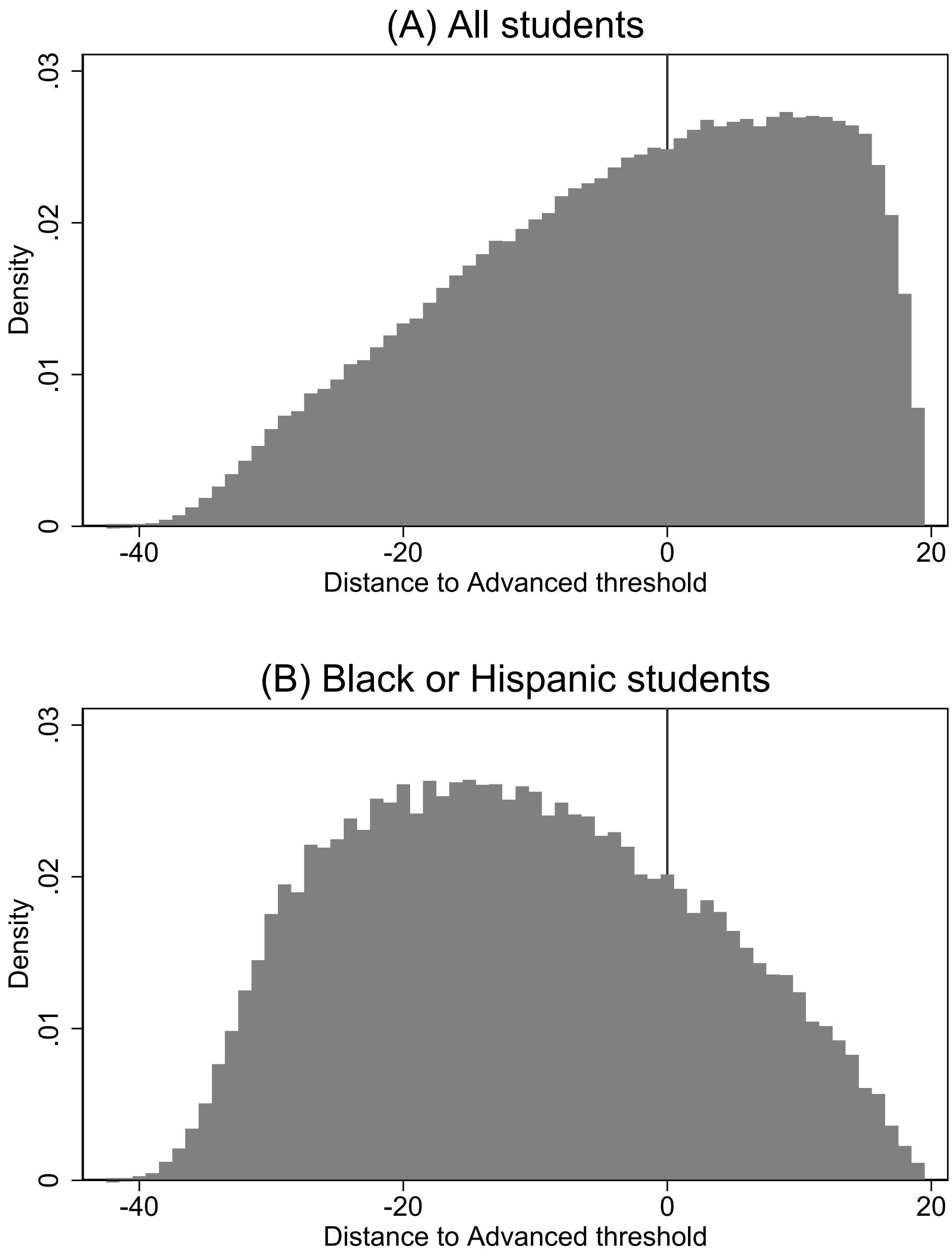
		English Language Arts	Mathematics	Science and Tech/Eng
Advanced Students at this level demonstrate a comprehensive and in-depth understanding of challenging subject matter and provide sophisticated solutions to complex problems.	High 270-280			
	Low 260-268			
Proficient Students at this level demonstrate a solid understanding of challenging subject matter and solve a wide variety of problems.	High 250-258			
	Low 240-248			
Needs Improvement Students at this level demonstrate a partial understanding of subject matter and solve some simple problems.	High 230-238			
	Low 220-228			
Warning/Failing Students at this level demonstrate a minimal understanding of subject matter and do not solve simple problems.	High 210-218			
	Low 200-208			
		2008 2009 **2010**	2008 2009 **2010**	**2010**

In the figure above, the top of the black bar indicates your child's score on each test. The small gray bar, included for 2010 only, shows the range of likely scores your child would receive if he or she took the test multiple times.

Example:
Your child's score → 240
← Range of likely scores if your child took the test many times

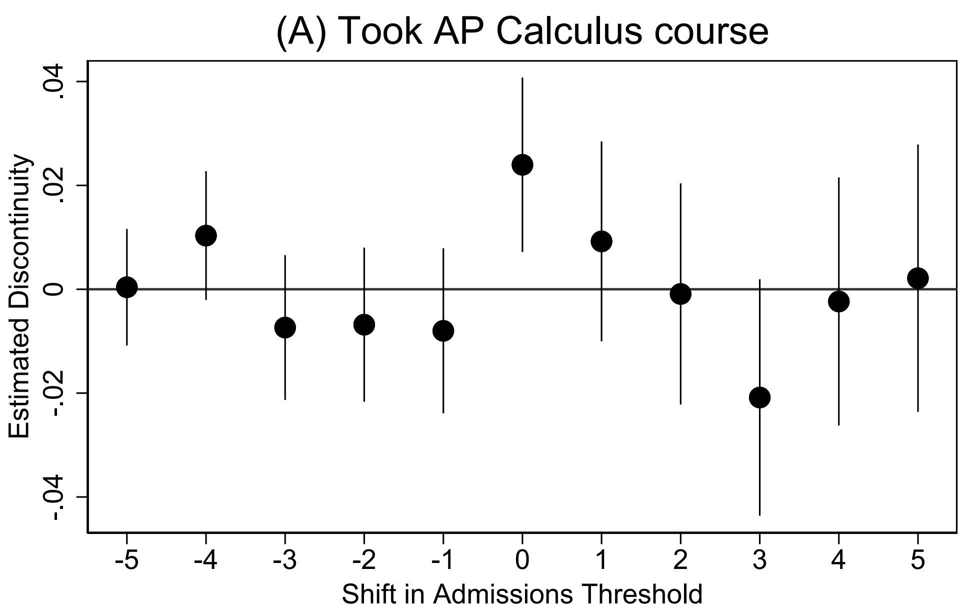
Notes: Shown above is the page where a student's MCAS results are first described to families in a report mailed to the student's home. This is the second page of the report, after a cover page explaining the purpose of the report. This report format remained stable throughout the cohorts studied in this paper.

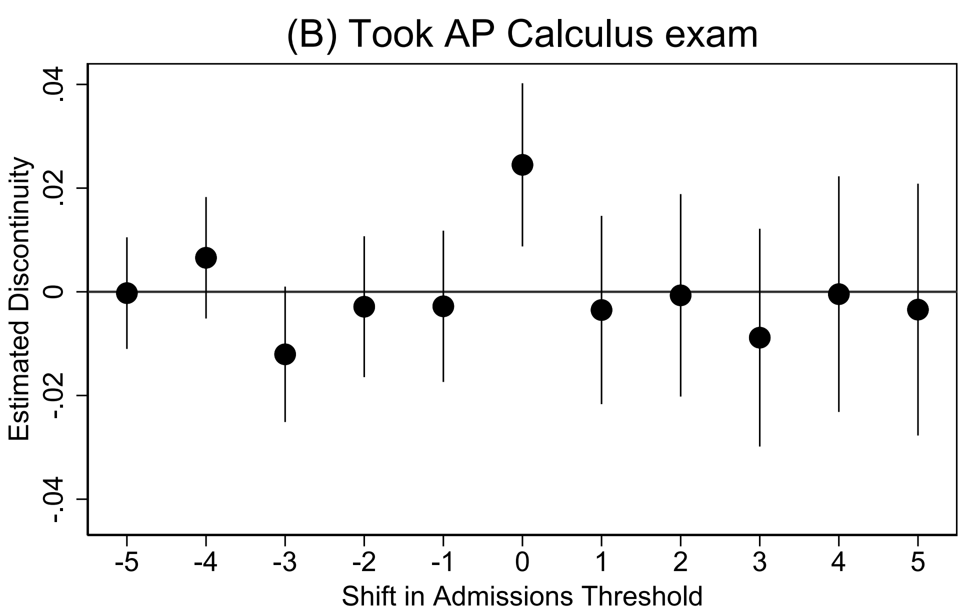
Figure A.2: Density of Running Variable



Notes: Shown above are histograms of the running variable for all students and for Black or Hispanic students.

Figure A.3: Placebo Tests: Black or Hispanic Students





Notes: Each estimate comes from a local linear regression discontinuity specification in which the threshold has been shifted by the listed number of raw score points. All regressions include high school by year fixed effects and use a bandwidth of four raw score points. Vertical lines indicate 95% confidence intervals. The sample includes all Black or Hispanic students.

Table A.1: Engagement and Success in Non-AP Calculus

				Black/Hispanic	
	All students (1)	White/ Asian (2)	Black/ Hispanic (3)	Low ELA scores (4)	High ELA scores (5)
Took non-AP Calculus course	-0.001	-0.002	0.001	-0.004	-0.004
	(0.004)	(0.004)	(0.009)	(0.013)	(0.013)
Control mean	0.069	0.070	0.072	0.072	0.071
A or B in non-AP Calculus course	0.001	0.001	-0.003	-0.008	-0.003
	(0.002)	(0.002)	(0.005)	(0.007)	(0.007)
Control mean	0.023	0.024	0.025	0.026	0.025
N	63,360	51,343	10,427	4,303	5,711

Notes: Heteroskedasticity robust standard errors clustered by high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Each estimate comes from a local linear regression discontinuity model that regresses an indicator for a given outcome on an indicator for receiving an score on the 10th grade math MCAS, raw score distance from that threshold, and the interaction of the two. All regressions include high school by graduating class fixed effects and use a bandwidth of four raw score points. The sample includes all students whose 10th grade MCAS timing implies they are in the high school classes of 2011 through 2014. Outcomes are defined as taking or earning at least a B in a Calculus course that is not Advanced Placement. Columns 4 and 5 limit the sample to Black or Hispanic students with ELA scores below or above the statewide average for that cohort. Listed below each column is the mean of the outcome for those one raw score point below the threshold.

Table A.2: Robustness Checks, Black or Hispanic Students

	Took AP Calculus course (1)	Took AP Calculus exam (2)
Bandwith = 4, linear	0.025***	0.024***
	(0.009)	(0.008)
+ Demographic controls	0.025***	0.024***
	(0.009)	(0.008)
+ Bandwidth = 8, quadratic	0.019**	0.021**
	(0.009)	(0.008)
+ Global RD, quadratic	0.027***	0.024***
	(0.006)	(0.006)

Notes: Heteroskedasticity robust standard errors clustered by distance to the threshold are in parentheses (* p<.10 ** p<.05 *** p<.01). The top row of each panel shows estimates from the baseline local linear regression discontinuity specification, which includes high school by year fixed effects and uses a bandwidth of four raw score points. The second row adds to the baseline specification controls for gender, race and low income status. The third row adds to the baseline specification quadratic terms in the running variable on either side of the threshold, using a bandwidth of eight raw score points. The fourth row uses that same quadratic specification estimated using all observations regardless of distance to the threshold. The sample is limited to Black or Hispanic students.

Table A.3: Placebo Outcomes and Running Variable

	All students (1)	White/ Asian (2)	Black/ Hispanic (3)
(A) ELA Advanced signal			
Took AP Calculus course	0.003	0.002	-0.000
	(0.004)	(0.005)	(0.010)
Control mean	0.131	0.133	0.147
Took AP Calculus exam	0.004	0.004	-0.001
	(0.004)	(0.005)	(0.010)
Control mean	0.123	0.125	0.141
N	112,317	96,889	12,891
(B) Non-math subjects			
Other AP exams taken	-0.002	-0.007	0.039
	(0.017)	(0.018)	(0.047)
Control mean	0.453	0.466	0.476
N	63,360	51,343	10,427

Notes: Heteroskedasticity robust standard errors clustered by high school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each estimate comes from a local linear regression discontinuity model that regresses an indicator for a given outcome on an indicator for receiving an score on the 10th grade math MCAS, raw score distance from that threshold, and the interaction of the two. All regressions include high school by graduating class fixed effects and use a bandwidth of four raw score points. The sample includes all students whose 10th grade MCAS timing implies they are in the high school classes of 2011 through 2014. Panel A uses 10th grade ELA scores as the running variable. Panel B uses 10th grade math scores as the running variable but uses as an outcome the total number of AP exams taken, excluding AP Calculus exams. Listed below each column is the mean of the outcome for those one raw score point below the threshold.

Suggested Citation: Avery C., Goodman, J. (2022). *Ability Signals and Rigorous Coursework: Evidence from AP Calculus Participation* (Working Paper 2022-#). Wheelock Educational Policy Center. Available at wheelockpolicycenter.org.

OUR MISSION

The Wheelock Educational Policy Center (WEPC) conducts and disseminates rigorous, policy-relevant education research in partnership with local, state, and federal policymakers and stakeholders to improve pk-20 educational opportunities and holistic outcomes for underserved students.

www.wheelockpolicycenter.org
wheelockpolicy@bu.edu

Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

WHEELOCK EDUCATIONAL POLICY CENTER