sociological science

Differences in Academic Preparedness Do Not Fully Explain Black–White Enrollment Disparities in Advanced High School Coursework

João M. Souto-Maior,[a] Ravi Shroff[a]

a) New York University

Abstract: Whether racial disparities in enrollment in advanced high school coursework can be attributed to differences in prior academic preparation is a central question in sociological research and education policy. However, previous investigations face methodological limitations, for they compare race-specific enrollment rates of students after adjusting for characteristics only partially related to their academic preparedness for advanced coursework. Informed by a recently-developed statistical technique, we propose and estimate a novel measure of students' academic preparedness and use administrative data from the New York City Department of Education to measure differences in Advanced Placement (AP) mathematics enrollment rates among similarly prepared students of different races. We find that preexisting differences in academic preparation do not fully explain the under-representation of Black students relative to White students in AP mathematics. Our results imply that achieving equal opportunities for AP enrollment not only requires equalizing earlier academic experiences, but also addressing inequities that emerge from coursework placement processes.

Keywords: racial disparities; organizational decisions; machine learning; high school coursework; academic preparedness

Replication Package: See the Data and Code Availability Statement on page 158.

Citation: Souto-Maior, João, and Ravi Shroff. 2024. "Differences in academic preparedness do not fully explain Black-White enrollment disparities in advanced high school coursework." Sociological Science 11: 138-163.

Received: October 30, 2023

Accepted: December 19, 2023

Published: March 11, 2024

Editor(s): Arnout van de Rijt, Jeremy Freese

DOI: 10.15195/v11.a6

The completion of advanced high school curricula—such as Advanced Placement (AP) and honors courses—influences many educational outcomes, including graduation rates, college enrollment, and college completion (Chajewski et al. 2011; Fischer 2007; Morgan and Klaric 2007; Engberg and Wolniak 2010). In the United States, the well-documented underrepresentation of Black students in advanced courses (Lucas et al. 2020; Xu et al. 2021; Riegle-Crumb et al. 2018), therefore, is concerning, as it contributes to the persistence of Black–White inequalities in educational experiences.

Attempts to mitigate Black–White gaps in advanced course-taking often focus on inequalities which arise from differences in the availability of advanced courses between schools (Roscigno et al. 2006; Iatarola et al. 2011), but equalizing the availability of advanced courses across schools is unlikely to eliminate racial gaps in advanced enrollment because course-taking inequalities also arise within schools (Tyson 2011; Oakes 2005; Mickelson 2001; Lewis and Diamond 2015; Xu et al. 2021; Klugman 2013).

A central challenge in crafting policies to reduce within-school racial disparities is to disentangle the extent to which advanced enrollment practices reproduce existing inequalities from the extent to which they exacerbate inequalities by favoring one or more racial groups (Morton and Riegle-Crumb 2019). That is, because advanced courses are, by definition, selective courses reserved for students with strong academic qualifications, observed racial disparities in advanced enrollment may largely reflect academic inequalities arising in earlier years of school (Malkus 2016; Morton and Riegle-Crumb 2019), with some suggesting, for example, that "the ultimate solution to AP participation gaps is closing the preparation gaps before high school" (Malkus 2016: p.11). In contrast, qualitative evidence indicates that advanced enrollment processes may favor similarly prepared White students over Black students (Carter 2005; Lewis and Diamond 2015; Oakes 2005; Tyson 2011)—exacerbating rather than simply reproducing disparities—suggesting that interventions should aim to reduce disparities stemming from the enrollment process itself rather than just equalizing academic preparation before high school.

To investigate the extent to which enrollment decisions tend to favor particular race groups over others, scholars often attempt to compare the course placement decisions made for students of different races but similar levels of *academic preparedness*. This question has received considerable attention in the literature, but previous studies come to markedly different conclusions (Gamoran and Mare 1989; Conger et al. 2009; Irizarry 2021; Kelly 2009; Lucas and Berends 2007). As we discuss in more detail in Online Supplement A, this approach stems from the fact that course placement decisions are based, at least ostensibly, on the principle that students should be assigned to the academic environment which best matches their prior academic experiences and current academic capabilities (College Board 2023a; Kelly 2007; Oakes and Guiton 1995; Kelly and Price 2011). However, the predominant quantitative methodology traditionally used to investigate this issue faces significant statistical shortcomings that limit its ability to accurately estimate enrollment disparities between students of different races with similar levels of academic preparedness for advanced coursework.

The Traditional Approach and Its Limitations

Traditionally, studies attempt to compare the course enrollment of different-race students with similar levels of academic preparedness by statistically adjusting for measured indicators of academic background in a regression model (Conger et al. 2009; Gamoran 1992; Kelly 2009). For example, studies often regress a binary indicator of students' advanced enrollment status on a race indicator and on several measures of students' academic history—for example, previous course grades, GPA, course-taking patterns, and scores on standardized exams.[1]

A nonzero (and statistically significant) estimated coefficient on the race indicator suggests—assuming the model is correctly specified—that students of different races with similar values of academic background measures enroll in advanced coursework at different rates. This is interpreted as evidence that equally prepared students of different races enroll at different rates, or equivalently, that academic preparedness does not fully explain racial disparities in advanced enrollment. Following this general strategy, studies have come to markedly different conclusions. Some studies suggest that various combinations of academic background variables fully account for the under-representation of Black students in advanced high school courses (Gamoran and Mare 1989; Hallinan 1992; Lucas and Gamoran 2002; Schiller and Hunt 2011; Conger et al. 2009), whereas others find that racial disparities persist even after adjusting for many relevant covariates (Gamoran 1992; Mickelson 2001; Lee and Bryk 1988; Irizarry 2021; Kelly 2009; Lucas and Berends 2007).

A central assumption underlying this approach is that students of different races are "similarly prepared" for advanced coursework when they have similar values on some set of measured indicators of academic background, like previous course grades and standardized exam scores. However, we argue that in practice, given a chosen set of measures of academic background, this assumption is problematic for two reasons. First, students who appear similar on the selected measures may in reality be differently prepared for advanced coursework, as students may differ on relevant measures that are unrecorded or unavailable; in a regression, this corresponds to the well-known issue of omitted-variable bias. Second, students who appear dissimilar on the selected measures may in reality be equally prepared for advanced coursework, for instance, if an included measure is unrelated or only partially related to a student's preparedness; this corresponds to the lesser-known challenge of included-variable bias (Ayres 2005).

To illustrate these issues, consider a school in which Black students enroll in AP Calculus at a lower rate than White students, and imagine that a researcher relies on the traditional regression approach described above to investigate whether differences in academic preparedness for AP Calculus explain this disparity. Suppose the researcher already adjusts for several measures of academic background in a regression of enrollment on race—for example, information on a student's prior mathematics coursework, prior mathematics GPA, and performance on standardized mathematics exams—and is faced with the option of including a student's previous performance in an 8th grade English class, a measure which may vary across race groups. Should the researcher adjust for this additional measure?

On one hand, failing to include 8th grade English performance in the model may result in *omitted-variable bias* if the measure provides additional information that is predictive of a student's preparedness for AP Calculus. If, for instance, 8th grade English performance is correlated with academic preparedness, conditional on the variables already included in the model, excluding this variable could skew estimates of the true role of academic preparedness in explaining the White advantage in AP Calculus enrollment. In fact, the notion that controlling for additional measures of academic background improves the reliability of empirical analyses investigating enrollment rate disparities is prominent in the literature, for example, studies that control for students' performance on standardized tests during middle school (Conger et al. 2009; Champion and Mesa 2016; Mickelson 2001) have been criticized for not including more complete measures of academic background such as course grades and course-taking trajectories (Champion and Mesa 2016; Irizarry 2021; Kelly 2009; Archbald and Farley-Ripple 2012).

On the other hand, adjusting for 8th grade English performance runs the risk of introducing *included variable bias* (Jung et al. 2019; Ayres 2005). If, conditional on covariates already included in the regression, 8th grade English performance has no (or only partial) correlation with preparedness for AP Calculus, then adjusting for this variable could again skew estimates of the true role of academic preparedness in explaining the White advantage in AP Calculus enrollment.

As illustrated by this example, the traditional approach to account for "academic preparedness" by adjusting for a set of student academic background measures, therefore, leads to uncertainty around precisely which measures of academic background one should include in a regression model. As evidenced by the many conflicting results in the literature, this uncertainty can lead to substantial variation in the measures of academic background used as controls and, ultimately, to unreliable estimates of advanced enrollment disparities between similarly prepared students of different races.

The Current Study

In this article, we address these methodological challenges by adapting a recently developed statistical approach from the discrimination literature (Jung et al. 2019) to analyze the role of academic preparedness in explaining the under-representation of Black students (relative to White students) in AP mathematics courses in New York City public schools. We estimate what we refer to as "preparedness-adjusted" enrollment disparities between racial groups, that is, disparities between students of different races who are similarly prepared for AP mathematics courses.[2]

Our approach has two novel aspects relative to the traditional approaches described above. First, rather than operationalizing academic preparedness via some chosen set of academic background measures, we define a student's preparedness for an advanced course with a single measure—their *ex-ante* probability of "success" in the course—and estimate this quantity. By directly adjusting for a student's estimated preparedness in a regression (along with a race indicator), we avoid the issue of included-variable bias. Second, we conduct a statistical sensitivity analysis to assess the robustness of our "preparedness-adjusted" estimates to the presence of plausible unmeasured confounding, mitigating the challenges posed by omitted-variable bias.

Using this approach, we analyze administrative data from more than 40,000 high school students in New York City and find that Black students have roughly 30 percent lower odds of enrollment in AP math, on average, compared to similarly prepared White students in the same school. Importantly, we contrast our results with those produced by traditional regression approaches and show that such approaches may underestimate the extent of preparedness-adjusted enrollment disparities between Black and White students.

Our study, therefore, provides two main contributions. First, by providing an estimate of racial disparities in AP coursework between similarly prepared students that is less vulnerable to the methodological limitations of previous analyses, our study furnishes evidence to inform interventions for mitigating persistent within-school racial gaps in AP course-taking. Second, by adapting a recently developed statistical technique to the context of course enrollment decisions, our study provides a more methodologically sound framework to address an important and well-known sociological inquiry: the comparison of course-taking patterns for students of different races but similar levels of academic preparedness.

Data and Measures

Our study focuses on participation in AP classes, a series of selective high school courses that provide qualified and academically motivated students with college-level coursework before they finish high school. The AP program is administered by the College Board and enables registered schools to provide AP courses in one of 38 different subjects (College Board 2023a). As part of the program, the College Board administers standardized AP exams to evaluate students' knowledge of AP subjects (Hacsi 2004). Colleges and universities generally allow students with sufficiently high AP exam scores to earn college credit or placement into specific courses, which provides students with a quicker transition into advanced courses during their college careers (Evans 2019). Our results focus on racial disparities in enrollment into a set of AP mathematics courses we refer to as "AP math," a choice which we discuss below.

We analyze a longitudinal, student-level data archive of all New York City public high schools, as provided by the New York City Department of Education. We focus on students who enrolled in 9th grade in 2011 or 2012 in public high schools in this school system, and who advanced exactly one grade each year between 7th and 12th grades. Our data, described in detail below (with additional details provided in Online Supplement B), include rich longitudinal information on students' academic background, allowing us to construct an extensive set of variables characterizing students' academic trajectories between 7th and 12th grades as well as their schools. Next, we discuss our measures of interest as well as the specific student- and school-level data restrictions we impose to create our sample.

AP Mathematics

We focus on student enrollment in a set of AP courses we refer to as "AP math," consisting of AP Calculus AB, AP Calculus BC, and AP Statistics. This grouping is based on the College Board's broad category of "AP Math and Computer Science," which includes the three courses just mentioned, along with AP Computer Science A and AP Computer Science Principles. We restrict our focus to calculus and statistics courses because of the limited availability of AP computer science courses in the high schools in our data, and the fact that only a small share of students who take AP calculus and statistics courses also take AP computer science courses.

We concentrate on AP mathematics courses for two reasons. First, these courses are highly influential on students' educational trajectories and downstream economic opportunities. Among the different AP subjects, studies show that the educational and economic benefits of STEM (science, technology, engineering and mathematics) courses stand out (Domina and Saldana 2012; Rose and Betts 2004). Further, given the changing dynamics of labor markets, an emphasis on STEM curricula is increasingly seen as important both for individual job prospects and national economic competitiveness (NSB (National Science Board) 2020). At the high school level, mathematics coursework is known to be an important factor that influences enrollment in subsequent STEM courses (Douglas and Attewell 2017) and STEM career aspirations (Warne et al. 2019).

Second, AP math courses are particularly well-suited for our examination of racial disparities in course enrollment. Our approach assumes that the specific measurement of students' course performance that we consider, performance on an associated AP exam, is not itself influenced by racial bias. Although the fact that AP exams are graded by a third party—that is, the College Board—and not the person who teaches the course can mitigate potential discrimination based on student identity[3], the grading of math exams in particular likely involves less subjective judgment than the grading of other subjects.

We analyze student enrollment in least one of the three distinct courses comprising "AP math" (i.e., AP Calculus AB, AP Calculus BC, and AP Statistics) instead of focusing on a single course, or on each course separately, for two reasons. First, it strikes us as more policy-relevant to measure preparedness-adjusted enrollment disparities in a given discipline rather than in a given course. That is, it seems reasonable to assume that it is beneficial for all qualified students to take *some* AP math course in high school, but perhaps more debatable that any particular AP math course is more important than any other. Second, course-taking patterns in our data for AP Calculus AB, AP Calculus BC, and AP Statistics are quite similar for Black and White students (see Online Supplement C). Moreover, the structure of our data does not allow us to differentiate AP math exam outcomes across these three kinds of AP math courses.

Sample

Here, we describe student- and school-level data restrictions we impose to create our sample. We focus specifically on high schools that met the following conditions for each academic year that started between the Fall of 2011 and the Fall of 2015: (1) the school was operational in the given academic year, (2) the school was identified as a "general academic" institution that enrolled students in grades 9 through 12, and (3) at least one student in each of the 2011 and 2012 cohorts enrolled in at least one AP math course at the school in the first four years of their high school careers. As our main comparison of interest involves enrollment disparities between Black and White students, we further restrict to schools that enrolled at least one White and one Black student from the 2011 and 2012 cohorts.

We restrict our analysis to students in the 2011 and 2012 cohorts who enrolled in 9th grade in one of the high schools described above, and, given the relevance of students' race to our analysis, only consider students that reported their race. We further restrict to students that followed a standard grade promotion trajectory, advancing exactly one grade each year from 7th to 12th grade. This restriction ensures that we analyze students following similar academic trajectories, and that the variables we include in our models represent the same academic year for each student. In Online Supplement B, we provide additional details about our sample and show that students who follow a standard grade promotion trajectory comprise close to 60 percent of students in each grade (Table S.B.1) and are much more likely to enroll in AP math classes than students who do not follow a standard grade trajectory (Table S.B.2).

Finally, we filter out students who were not enrolled in one of the high schools considered above by the end of 12th grade, perhaps because they dropped out of school or transferred outside the public school system. We note that our sample includes students who transferred between high schools, as long as such transfers occurred within the high schools selected above (such transfers are rare). Figure S.B.1 in Online Supplement B details the racial composition of our sample, where each student's race is determined from 9th grade administrative records. After these restrictions, we are left with 42,469 students distributed across 115 high schools. In Online Supplement B, we show that with these student-level restrictions, our sample consists of the majority of AP math takers and AP math exam takers in the student population of interest.

Methods

Our overall goal is to compare AP math enrollment rates between students of different races with similar levels of "academic preparedness" for such courses. As noted above, traditional approaches which estimate a regression model of enrollment against race and various academic background measures—*implicitly* adjusting for student preparedness—run into the challenges of omitted- and included-variable bias. Our approach, in contrast, first specifies a novel measure of academic preparedness; we then estimate this quantity for every student in our sample and *explicitly* adjust for this measure in a regression of enrollment against student race.

Defining Academic Preparedness

At a high level, we propose that a student's academic preparedness for a given advanced course can be defined as their *ex-ante* chance of "success" if they were to take the course; we emphasize that this concept is defined *before* enrollment in the course occurs. We introduce the notion of "success" because the general rhetoric for sorting students across different-level courses based on academic preparedness relies on a concern that students might not succeed in academic environments for which they are not adequately prepared (Goldsmith 2011; Fitzpatrick and Mustillo 2020; Hallinan 1994; Lavy et al. 2012).[4] From this general rhetoric, it follows that students are academically prepared to take a given course if they are are expected to "succeed" in it (Oakes 2005; Hallinan 1994; Oakes and Guiton 1995; Fitzpatrick and Mustillo 2020; Lucas et al. 2020).

In the context of enrollment into AP math courses, we posit that "success" can be defined by whether or not the student would pass the corresponding AP exam, if they were to take the course and the exam. We focus on AP exams for two main reasons. First, AP courses function as preparation for AP exams and, thus, students' short-term purpose for taking AP courses is to pass the respective AP exams (Judson et al. 2019)—in fact, the relevance of AP exams is such that the educational benefits associated with AP course-taking are often conditional on students' scores on AP exams (Ackerman et al. 2013). Second, AP exams are administered and graded by the College Board, not the person who teaches the corresponding course (Hacsi 2004), mitigating potential discrimination based on student identity.

In our data, the vast majority of students who take any AP math courses do so in 11$^{\text{th}}$ or 12$^{\text{th}}$ grade (Table S.B.2). Therefore, we assume that enrollment processes begin at the start of 11$^{\text{th}}$ grade.[5] Then, our definition of academic preparedness involves three student-level variables of interest: *AP math enrollment*, indicating whether a student enrolled in at least one AP math course during 11$^{\text{th}}$ or 12$^{\text{th}}$ grade; *AP math exam participation*, indicating whether a student took at least one AP math exam during 11$^{\text{th}}$ or 12$^{\text{th}}$ grade; and *AP math exam passage*, indicating whether a student passed at least one AP math exam in 11$^{\text{th}}$ or 12$^{\text{th}}$ grade.[6] Following these measures, we define each student's *academic preparedness* to be the *ex-ante* probability—estimated at the start of 11$^{\text{th}}$ grade—that they would pass at least one AP math exam in grades 11 or 12, if they were to take at least one AP math course and at least one AP math exam. We reflect on this operationalization in the Discussion, where we show that alternative operationalizations do not have a substantial impact on our results.

More formally, we assume we have data of the form $\Omega = \{(c_i, a_i, t_i, r_i, x_i)\}_{i=1}^{N}$ for a collection of N students, where for the i^{th} student, c_i indicates their race, a_i is a binary indicator of whether they were assigned into at least one AP math course in grades 11 and 12 ($a_i = 1$ if the student enrolled in an AP math course and 0 otherwise), t_i is a binary indicator of whether they attempted at least one AP math exam in grades 11 and 12 ($t_i = 1$ if the student attempted an exam and 0 otherwise), r_i is a binary indicator of whether they passed at least one exam in grades 11 and 12 ($r_i = 1$ if the student passed an exam and 0 otherwise), and x_i indicates all other student covariates, measured before enrollment decisions occur, that is, before the start of grade 11. In what follows, we will refer to, for example, "passing an exam" as shorthand for passing at least one exam in grades 11 and 12.

Because whether a student passes an AP math exam may depend on whether they actually enroll in the course and take the exam, each student has four potential outcomes for exam passage (Imbens and Rubin 2015). We denote these potential outcomes by $r_i(a, t)$ for $a \in \{0, 1\}$ and $t \in \{0, 1\}$, where $r_i(a, t) = 1$ if the student would have passed the exam under enrollment condition a and exam-taking condition t.[7] Note that $r(1, 1)$ is only observed for students who actually enroll in an AP math course and take the exam—for these students, we can replace potential outcomes with the observed exam passage outcomes, setting $r(1, 1) = r$. With this setup, we then formally define a student's academic preparedness for an AP math course by

$$\mu_i = \Pr(r_i(1, 1) = 1 \mid c_i, x_i), \tag{1}$$

the probability the student would pass an exam if they were to both enroll in a course and take an exam.

Given this definition of academic preparedness, our approach for measuring preparedness-adjusted racial disparities in AP math consists of the following three steps, adapting the approach described in Jung et al. (2019) to our setting.

Step 1. Estimating Academic Preparedness

Using detailed longitudinal data, we fit a flexible machine learning model to estimate, for each student, the *ex-ante* probability that they would pass an AP math

Table 1: Complete and incomplete information students in our sample.

Group	N	% of sample
Complete information students		
AP math enrollment = yes; Exam participation = yes	8957	21.1%
Incomplete information students		
AP math enrollment = yes; Exam participation = no	1197	2.8%
AP math enrollment = no; Exam participation = no	31542	74.3%
AP math enrollment = no; Exam participation = yes	773	1.8%

exam if they were to both enroll in an AP math course and take an AP math exam in grades 11 or 12 (μ_i). It is challenging to estimate this probability accurately for every student, as not every student enrolls in AP math courses and not every student who enrolls in the advanced course takes AP math exams. That is, in the data, we only have full information on exam passage outcomes for students who, in reality, took AP math courses and AP math exams—we refer to these students as *complete information* students, and all other students as *incomplete information* students. If we assume, however, that conditional on observed covariates, exam passage potential outcomes are independent of actual AP enrollment and AP exam-taking decisions (i.e., that both enrollment and exam-taking decisions are *conditionally ignorable*), then we can estimate academic preparedness accurately for each student using observed data (Jung et al. 2019).

Formally, if we assume:

$$r_i(1,1) \perp (a_i, t_i) \mid c_i, x_i, \tag{2}$$

then we can express preparedness as follows:

$$\begin{aligned} \mu_i &= \Pr(r_i(1,1) = 1 \mid c_i, x_i) \\ &= \Pr(r_i(1,1) = 1 \mid t_i = 1, a_i = 1, c_i, x_i) \\ &= \Pr(r_i = 1 \mid t_i = 1, a_i = 1, c_i, x_i), \end{aligned} \tag{3}$$

where we can estimate this last quantity from observed data. In the second equality above, we used the conditional ignorability assumption (Eq. 2), and in the third equality we replaced potential outcomes with observed exam passage outcomes.

To estimate preparedness, we first restrict our data to the 20 percent of students who enrolled in at least one AP math course and took at least one AP math exam (i.e., the complete information students, Table 1).

Next, we train an XGBoost extreme decision trees model (Chen and Guestrin 2016) on a random subset of 90 percent of complete information students, estimating the probability of passing at least one AP math exam as a function of an extensive set of student- and school-level covariates measured before the start of 11$^{\text{th}}$ grade; this is our "exam passage model."[8] Table S.E.1 in Online Supplement E details all the covariates included in this model. We emphasize that this strategy circumvents the challenge of deciding exactly which variables might encode a student's academic preparedness (the motivating challenge behind the approach presented in this article). In our exam passage model, we include *all measures available in our data* (except student race; see below); in effect, the model is responsible for determining the extent to which covariates predict success in AP math courses.

This estimated exam passage model has an out-of-sample AUC of 0.877 (estimated on the remaining held-out 10 percent of complete information students), and appears well-calibrated across race groups (Figure S.E.1 in Online Supplement E). We do not use a student's race as a predictor of their academic preparedness, because an improvement in the predictive performance of the exam passage model could be due to well-known Black–White differences in learning experiences after enrollment in advanced courses (Lubienski 2002; Hallett and Venegas 2011; Riegle-Crumb and Grodsky 2010). That is, if estimates of academic potential are influenced by predicted future barriers to learning, such estimates no longer capture current academic preparedness.

Finally, we use the trained exam passage model to estimate our measure of academic preparedness—the *ex-ante* probability of passing at least one AP math exam (if one were to take an AP math course and an AP math exam)—for every student in our sample.

The reliability of our model's extrapolation from complete information students to incomplete information students rests on the assumption that no unmeasured variables confound student enrollment and exam-taking decisions and exam passage potential outcomes; we relax this conditional ignorability assumption in our sensitivity analysis below.

Step 2. Estimating preparedness-adjusted enrollment disparities

Given the estimates of academic preparedness for each student calculated in Step 1, we estimate preparedness-adjusted disparities across race groups via a logistic regression model (Jung et al. 2019), which we call a "preparedness-adjusted" regression.

This statistical model estimates the probability that a student enrolls in the advanced course, adjusting for student race, academic preparedness, and—because advanced enrollment might depend on school-level characteristics (Lucas and Berends 2007; Kelly 2009)—an indicator for the school the student attends.

Formally, this "preparedness-adjusted regression" can be written as follows:

$$\Pr(a_i = 1 \mid c_i, \mu_i, s_i) = \text{logit}^{-1}(\theta_0 + \theta_{c_i} + \theta_1 \text{logit}(\mu_i) + \theta_{s_i}), \tag{4}$$

where for the i^{th} student with race c_i in school s_i, we have $a_i = 1$ if the student enrolled in the advanced course and 0 otherwise; θ_{c_i} is the coefficient for the student's race (the main quantity of interest); μ_i is the student's academic preparedness; and θ_{s_i} is a school fixed effect. If, for example, White students are defined as the reference category, then $\hat{\theta}_{\text{Black}}$ is our main quantity of interest. If $\hat{\theta}_{\text{Black}}$ were negative, it would suggest that Black students were less likely to enroll in the advanced course than similarly prepared White students in the same school. In Eq. 4, we apply a logit transformation to μ_i to reflect an assumption that the log-odds of enrolling in the advanced course are approximately proportional to the log-odds of preparedness, although other transformations may make sense, depending on the context (Jung et al. 2019).

Step 3. Assessing robustness of estimates to violations of ignorability

The assumption that no unmeasured variables confound enrollment and exam-taking decisions and exam passage potential outcomes (i.e., the conditional ignorability assumption from Eq. 2) allows us to accurately estimate academic preparedness for all students, and hence to estimate preparedness-adjusted disparities using Eq. 4. However, this assumption is unlikely to hold exactly in our context, as students who actually enroll in AP math and take an AP math exam may differ from students who do not enroll in AP math (or who enroll but do not take an exam) in ways that are not recorded in our data. If such unrecorded factors influence exam passage, our preparedness-adjusted approach may produce skewed estimates of θ_c.

To assess the sensitivity of estimates of preparedness-adjusted disparities from Eq. 4 to possible unmeasured confounding, we adapt the approach of Rosenbaum and Rubin (1983). Specifically, we assume that there exists an unmeasured binary variable u that affects both student enrollment decisions and their chances of passing an exam, if they were to take the course and the exam. We also assume that among enrolled students, u does not affect the decision to take the exam, although this assumption could be relaxed in a more detailed sensitivity analysis. Finally, we assume that u accounts for all such confounding: if we knew its value for each student along with the other measured covariates, we could then accurately estimate preparedness for all students in our sample. Formally, our assumptions for the sensitivity analysis can be written as follows:

$$r(1,1) \perp (a,t) \mid c, x, u \tag{5}$$

$$u \perp t \mid a = 1, c, x. \tag{6}$$

Following Rosenbaum and Rubin (1983), we assume that the following parameters governing the prevalence and nature of the unmeasured confounder u are specified:

1. $q_{c,x} = \Pr(u = 1 \mid c, x)$, the prevalence of the confounder u;
2. $\alpha_{c,x}$, the effect of u on enrollment decisions; and
3. $\delta_{c,x}$, the effect of u on the exam passage outcome, conditional on taking a course and an exam.

As the notation indicates, the parameters $(q_{c,x}, \alpha_{c,x}, \delta_{c,x})$ can be specified separately within covariate strata (c, x). Then, given the values of the three parameters defined above, we can derive preparedness-adjusted estimates that account for the confounder u; details of the derivation are provided in Online Supplement D.

Because, by definition, one does not usually know characteristics of an unmeasured confounder, we derive preparedness-adjusted estimates that account for u by searching over a grid of plausible ranges for parameters $q_{c,x}$, $\alpha_{c,x}$, and $\delta_{c,x}$.

For each parameter combination, we obtain a single preparedness-adjusted estimate that accounts for u. We report the largest and smallest resulting estimates of

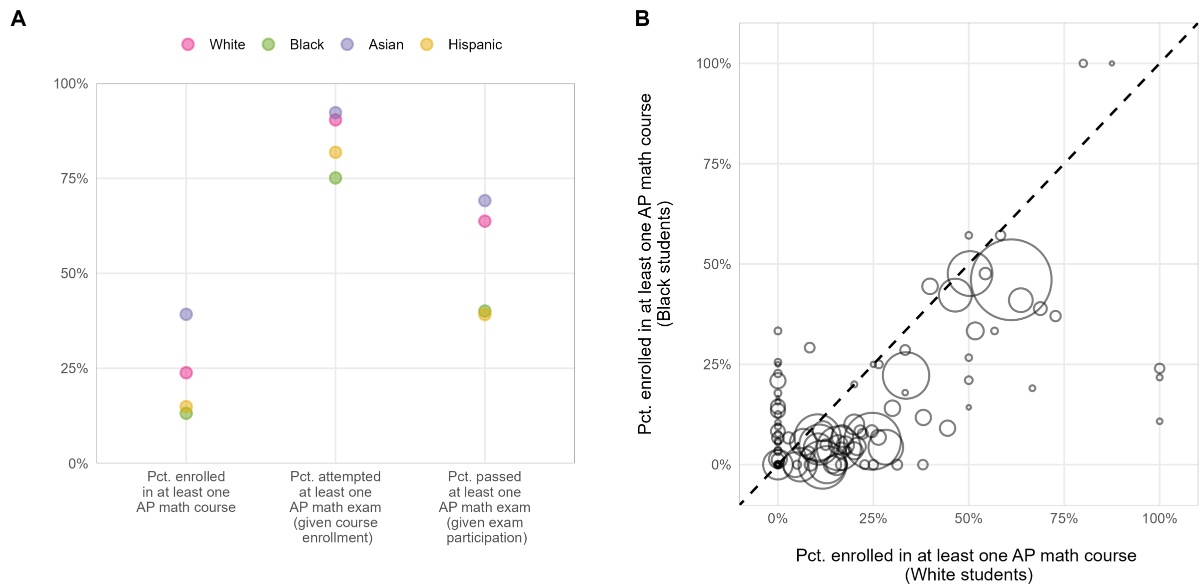

Figure 1: Descriptive analysis of racial disparities in AP math participation.

(a) AP math participation by race group for students in our sample. There are large racial disparities in AP math enrollment and AP exam passage rates among those who took an exam. However, AP exam participation is more homogeneous across racial groups.

(b) Black–White disparities in AP math enrollment in our sample, disaggregated by school. Each point represents a high school and is sized by the number of students in that school who are part of our sample. Among almost all schools that have some Black and White AP math enrollment, enrollment rates are higher for White students than for Black students.

preparedness-adjusted disparities, accounting for confounding. For computational feasibility, we assume that $\alpha_{c,x} = \alpha$ and $\delta_{c,x} = \delta$ (i.e., that these parameters are constant across all students) with $\alpha, \delta \in [-\Theta, \Theta]$ for some specified $\Theta > 0$, and that $q_{c,x} = q_c$ only depends on student race.

Results

Descriptive Analysis of Racial Disparities

Figure 1a summarizes raw disparities in AP math enrollment, exam participation, and exam passage for White, Black, Hispanic, and Asian students in our sample. Although our focus in this article is on Black–White disparities, we present descriptive statistics for other ethnoracial groups for context. Students whose race was reported as ‘other’ are omitted from the figure; they represent about 8 percent of the sample.

Although roughly 39 percent of Asian students and 24 percent of White students enrolled in at least one AP math course by the end of 12$^{\text{th}}$ grade, only about 13 percent of Black students and 15 percent of Hispanic students did the same.

Among those who enrolled in AP math, students took AP math exams at high rates, although differences between racial groups remain. Finally, among those who took an AP exam, Black and Hispanic students were much less likely to pass than White and Asian students: 40 percent of Black students and 39 percent of Hispanic students passed at least one AP math exam, whereas 64 percent of White students and 69 percent of Asian students passed an exam.

Because AP math course-taking patterns vary across schools—due, in part, to differences in both the availability of such courses and how schools structure their enrollment processes (Iatarola et al. 2011; Kelly 2007)—we display AP math enrollment rates for Black and White students within each school in our sample in Figure 1b. For almost all schools that have some Black AP math enrollment and some White AP math enrollment, enrollment rates are higher for White students than for Black students. Overall, our data demonstrate large enrollment disparities between Black and White students, consistent with known national patterns (Xu et al. 2021). Our goal is to measure the extent to which such disparities are explained by differences in student academic preparedness, hence we next estimate the preparedness of each student in our sample.

Step 1. Estimating Academic Preparedness

The first step in our approach is to apply the trained exam passage model (described in the Methods section) to estimate our measure of academic preparedness—that is, the *ex-ante* probability of passing at least one AP math exam, if one were to take an AP math course and an AP math exam—for every student in our sample. Figure 2 displays the distribution of estimated *ex-ante* probabilities of AP math success across race groups.[9] We observe that, on average, Asian students tend to be the most prepared, followed by White students, and then Hispanic and Black students, suggesting that some of the AP math enrollment disparities observed in Figure 1a may be explained by group differences in academic preparedness.

As the reliability of these estimates for *incomplete information students* rests on the assumption that no unmeasured variables confound student enrollment (and exam-taking) decisions and exam passage potential outcomes, we complement our results with a sensitivity analysis (Step 3, below), where we relax this assumption.

Step 2. Estimating Preparedness-Adjusted Enrollment Disparities

Next, we fit our preparedness-adjusted regression (Eq. 4) to gauge the extent of enrollment disparities in AP math between similarly prepared students from different race groups. Our estimates of preparedness-adjusted disparities are presented in the left panel of Figure 3 (labeled "Preparedness-adjusted"). Because our estimates of academic preparedness depend on our sample, we compute the confidence intervals for coefficients displayed in this panel via bootstrapping.[10] Estimated coefficients from all models presented in Figure 3 are available in Table S.E.2 in Online Supplement E. As in Figure 1a, although our focus is on Black–White enrollment disparities, we provide estimated disparities between other ethnoracial groups and White students for context. We find that Black and Hispanic students have, on average, roughly 30 percent lower odds of enrollment in at least one AP math course

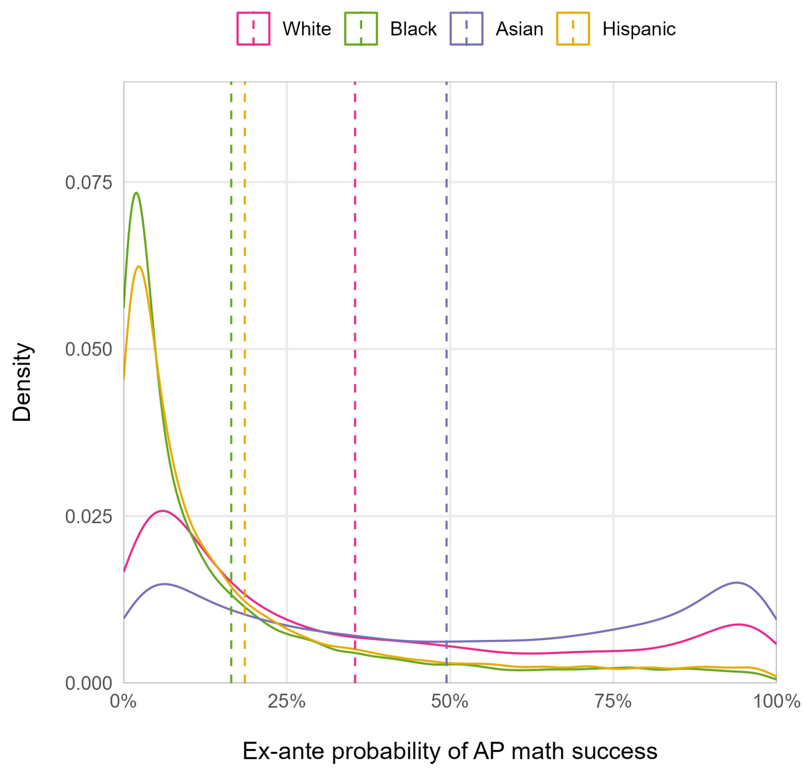

Figure 2: The distribution of estimated *ex-ante* probability of AP math success across race groups. Results are presented for all students, using the exam passage model trained on students that took at least one AP math course and at least one AP math exam. The mean of each distribution is indicated with a dashed vertical line. AP math success is defined as passing at least one AP math exam if one were to take at least one AP math course and at least one AP math exam. The distributions indicate that by the start of 11th grade, White and Asian students are, on average, more academically prepared to take AP math courses than Black and Hispanic students.

compared to similarly prepared White students in the same school. Asian students, on the other hand, have about 60 percent higher odds of AP math enrollment than similarly prepared White students in the same school. In particular, low enrollment rates for Black students in AP math courses relative to White students are unlikely to be fully explained by preexisting differences in preparedness for those courses.

We contrast our results with several alternative approaches: unadjusted within-school disparities and two traditional regression-based approaches similar to those in the literature; model specifications and fitted coefficients are available in Table S.E.2 in Online Supplement E.

In the second panel in Figure 3 (labeled "Raw disparities"), we display average within-school racial disparities in AP math enrollment between non-White students and White students estimated via a logistic regression of AP math enrollment on race, adjusting only for the school attended. On average, Black and Hispanic

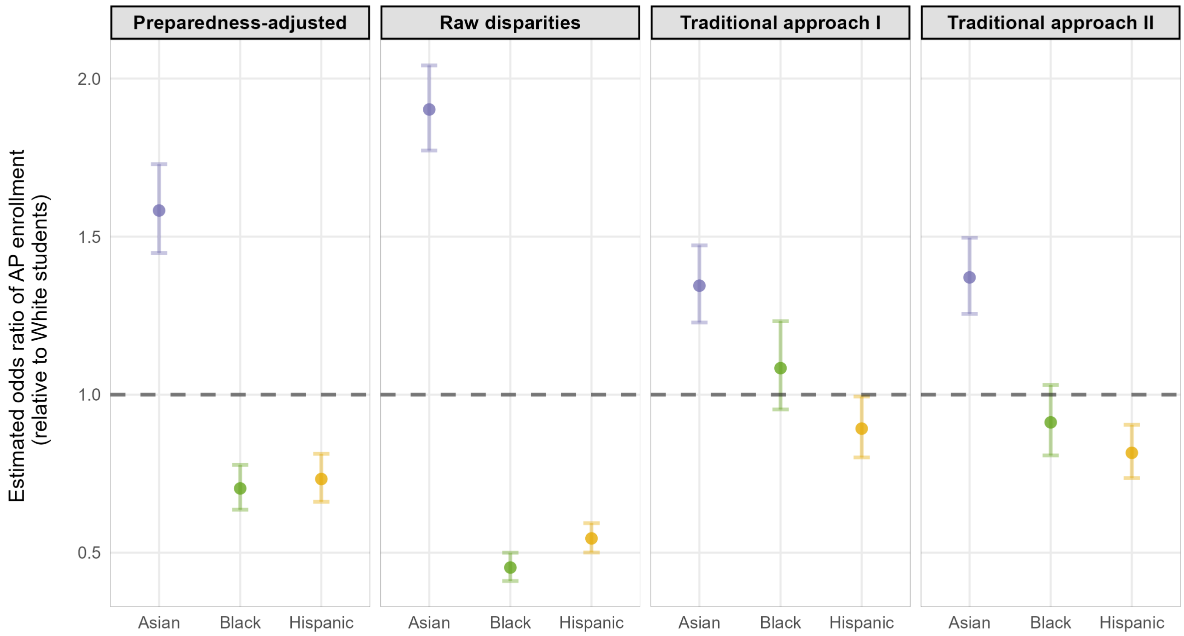

Figure 3: Adjusted disparities in AP math enrollment measured four ways. Each approach consists of a logistic regression estimating enrollment in AP math as a function of race and other covariates. The preparedness-adjusted model adjusts for school attended and estimated academic preparedness, the raw disparities model adjusts only for school, the traditional approach I model adjusts for school and an extensive set of student characteristics, and the traditional approach II model adjusts for school and a smaller selected set of student characteristics. Each panel presents odds ratios for enrollment of non-White students compared to White students; 95 percent confidence intervals for the preparedness-adjusted model are computed via bootstrapping. The raw disparities model suffers from omitted-variable bias, overestimating Black–White disparities by failing to adjust for variables related to academic preparedness. Both traditional approaches also suffer from included-variable bias by adjusting for variables only partially related to academic preparedness, and likely underestimate preparedness-adjusted Black–White disparities. The estimates from the preparedness-adjusted model, our preferred approach, demonstrate enrollment advantages of White students over similarly prepared Black students.

students have about half the odds of enrollment in AP math as White students in the same school, whereas Asian students have almost twice the odds of AP math enrollment as White students. These estimates, however, are an imperfect measure of preparedness-adjusted disparities in AP math enrollment, as they do not account for any preexisting differences in academic preparedness across race groups. For example, because Black students in our sample appear to be, on average, less prepared for AP math than White students (Figure 2), such raw disparities likely overestimate the extent of differential enrollment attributable to the enrollment process itself.

Next, we estimate a logistic regression of AP math enrollment that adjusts for an extensive set of student-level academic and social variables (including race and school attended) and present estimated odds ratios of non-White to White enrollment in the panel in Figure 3 labeled “Traditional approach I.” This approach aims to mitigate potential omitted-variable bias by comparing enrollment rates among students of different races who are similar along a wide range of characteristics. This model finds no significant enrollment disparities between similar Black and White students, but does find a statistically significant Asian advantage over similar White students, and a statistically significant White advantage over similar Hispanic students. Because this model directly adjusts for covariates that may only be tenuously related to success in AP math (e.g., GPA in English courses), the resulting estimates likely suffer from included-variable bias, underestimating the extent of preparedness-adjusted disparities. In contrast, by distilling all available covariates into an estimate of academic preparedness, our preparedness-adjusted regression only accounts for covariates to the extent that they actually predict success in AP math courses.

Finally, we estimate a logistic regression of AP math enrollment that attempts to strike a balance between included- and omitted-variable bias by only adjusting for selected academic and socioeconomic student characteristics, in addition to race and school attended (“Traditional approach II” in Figure 3). This model only adjusts for academic variables related to prior mathematics courses in high school, under the assumption that variables related to middle school coursework and to, for example, English and science courses, might introduce some level of included variable bias. The results from this model again differ from our preparedness-adjusted estimates, suggesting that simply controlling for some particular set of covariates is unlikely to accurately capture students’ academic preparedness. That is, the “Traditional approach II” model is both subject to omitted-variable bias by excluding information somewhat relevant to AP math preparedness (encoded in previous grades in non-math courses, for instance), while potentially still suffering from included-variable bias to the extent that the selected academic and sociodemographic characteristics are unrelated to AP math preparedness.

Step 3. Assessing Robustness of Estimates to Violations of Ignorability

The accuracy of the preparedness-adjusted estimates of enrollment disparities across race groups in Figure 3 relies on the assumption that no unmeasured variables confound the relationship between AP math enrollment, exam-taking, and AP math exam passage. To assess the robustness of this result to violations of this important assumption, we apply our sensitivity analysis approach described in the Methods section. Our sensitivity analysis involves a grid search over parameters defining the prevalence of an assumed unmeasured confounder u, the influence of u on AP math enrollment, and the influence of u on exam passage potential outcomes.

To determine the range for the grid search over α and δ (i.e., to determine Θ), we calibrate the influence of u to an estimate of the association between a known covariate, previous math GPA, and enrollment and exam passage. In particular, we binarize 10^{th} grade math GPA by coding grades higher than one standard deviation above the mean as 1, and other scores as 0. We then fit a logistic regression on our sample, predicting enrollment as a function of school, 9^{th} and 10^{th} grade high school English and science GPA, and the binarized math GPA variable, and find that students with a 10^{th} grade math GPA that is one standard deviation above the mean have almost triple the odds of AP math course enrollment than students with lower math GPA scores. Similarly, we fit a logistic regression on complete information students, predicting exam passage as a function of the same variables, and observe that students with a 10^{th} grade math GPA that is one standard deviation above the mean again have almost triple the odds of AP math exam passage compared to students with lower math GPA scores. This analysis suggests that even a variable like prior math GPA known to be predictive of both enrollment and exam passage is unlikely to triple the odds of AP enrollment/exam passage. Thus, we assume that u can at most triple the odds of enrollment (or divide them by three), and at most triple the odds of exam passage, conditional on taking an AP math course and AP math exam (or divide them by three)—that is, we set $\Theta = \log(3)$.[11]

In Figure 4, we display the maximum and minimum values of preparedness-adjusted disparities resulting from this grid search, along with 95 percent confidence bounds, to give a sense of how much our preparedness-adjusted estimates from Figure 3 could vary in response to unmeasured confounding of the specified type and extent. This sensitivity analysis indicates that our finding that Black students enroll in AP math courses at lower rates than similarly prepared White students is robust to substantial confounding.

Discussion

Advanced course-taking in high school is an important part of students' educational experiences in the United States. In this article, we examined the extent to which the enrollment process itself shapes the substantial disparities between Black and White students in AP math course-taking. Evidence on this issue has direct consequences for policy interventions to help equalize opportunities for advanced enrollment. If the placement process largely reproduces preexisting academic disparities, interventions should address inequalities in early-childhood and middle school education, and in the first two years of high school. However, if the placement process exacerbates existing Black–White academic disparities, such interventions should be coupled with modifications to the placement process.

To address this question, we analyzed preparedness-adjusted placement disparities in the AP math course-taking patterns of students in New York City public high schools. We discovered significant disparities in Black–White enrollment rates among similarly prepared students—that is, among students who had similar estimated chances of passing an AP math exam, were they to take an exam and course—and our findings are robust to a substantial degree of unmeasured confounding. Our analysis differs from traditional regression-based approaches which operationalize academic preparedness through a series of academic background measures. We argue that, in contrast to the traditional approach which dominates the literature, the approach we detail is less vulnerable to issues of omitted- and included-variable bias, and provides more reliable evidence on the extent to which differences in academic preparedness can explain disparities in course-taking.

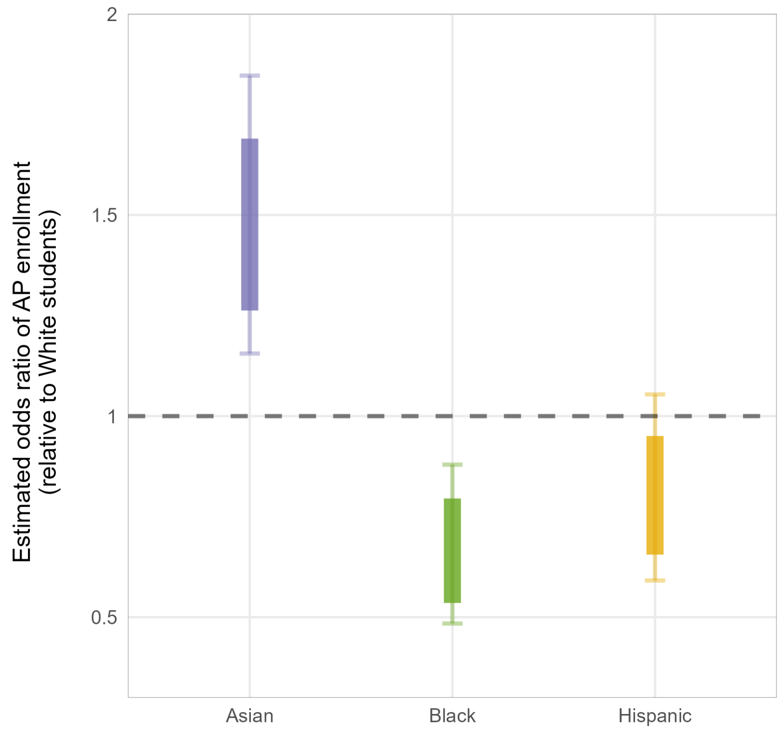

Figure 4: Assessing the sensitivity of preparedness-adjusted regression estimates to unmeasured confounding. We display estimates of preparedness-adjusted AP math enrollment disparities across race groups (relative to White students), accounting for a binary unmeasured confounding variable. The thick lines display sensitivity bands calculated by a grid search over parameters controlling the influence of unmeasured confounding. We assume that the confounder can at most triple the odds of enrollment (or divide them by three), and at most triple the odds of exam passage, conditional on taking an AP math course and an exam (or divide them by three). Confidence intervals on the ends of the sensitivity bands are formed by adding and subtracting 1.96 times the bootstrapped standard errors calculated previously.

We conclude by discussing limitations of our analysis, suggestions for future research, and policy recommendations.

Limitations and Future Directions

Although the strategy we propose in this article (adapted from (Jung et al. 2019)) addresses statistical limitations of common regression-based approaches, it is not without its own limitations. One challenge is that our approach may not produce accurate estimates of academic preparedness for students who never enrolled in the advanced course, or who enrolled and did not take an exam. Although we provide a detailed sensitivity analysis to address this issue, we note that, by definition, we can never know the structure and influence of factors that are unmeasured.

Another potential issue stems from the fact that Black and White students may have different learning experiences in advanced courses (Lubienski 2002; Hallett and Venegas 2011; Riegle-Crumb and Grodsky 2010). In our approach we assume that, among students with the same observed preenrollment covariates, all students' AP exam passage chances would be the same if they were to take the course and exam. However, if Black students perform worse on AP exams compared to White students in part because of racialized learning experiences within the classroom, rather than differences in academic preparation that exist before enrollment, our exam passage model would underestimate the preparedness of Black students, and our preparedness-adjusted regression would likely underestimate the extent of real preparedness-adjusted enrollment disparities. That is, our measure of academic preparedness estimated *before* enrollment may be biased by the fact that some students suffer discrimination in their learning experiences *after* enrollment.

Several factors mitigate this concern in our particular setting. First, although learning experiences within advanced courses may be racialized, AP exams are written and graded by a third party, the College Board. Therefore, biases present within the classroom might have a smaller effect on students' AP exam performance than they would on, for example, students' AP course performance. Second, student academic performance is influenced by a large set of factors, including student experiences in early years of schooling (Benson and Borman 2010; Clotfelter et al. 2009; Fryer and Levitt 2006) and out-of-school circumstances (Condron 2009; Chetty et al. 2020; Ainsworth 2002). Thus, the influence of racialized experiences in any particular class may have a relatively modest effect on exam passage. Most importantly, though, the policy implications of our study remain unchanged if we are in fact underestimating the extent of real preparedness-adjusted placement disparities. Therefore, this potential limitation does not reduce the significance of the practical conclusion that the placement process contributes to the exacerbation of Black–White academic inequalities in AP math enrollment.

Another potential concern is that our definition of academic preparedness presented — a student's probability of passing *at least one* AP math exam if they were to take at least one exam and at least one course—may not reflect students' actual preparedness if some students take more exams than others, and thus have more chances to pass at least one exam. This possible complication is unlikely to alter our substantive findings for two reasons. First, the distribution of exam-taking is similar across races: most students who take AP math courses tend to take one AP math exam, although White students are somewhat more likely to take two or more exams than Black students (see Figure S.C.2 in Online Supplement C). Second, we repeat our analysis with an alternative definition of academic preparedness, defined to be a student's *ex-ante* probability of passing *every* AP math exam they take (if they were to take at least one course and at least one exam). Instead of possibly overestimating the academic preparedness of students who take multiple exams, we may now underestimate the academic preparedness of such students. Rerunning our full analysis with this new definition of preparedness (i.e., estimating preparedness, fitting the preparedness-adjusted regression, and performing the sensitivity analysis) yields virtually identical findings to those presented above. This is because less than 5 percent of students in our data who took an AP course and an AP exam actually have a different “success” status under the alternative definition. Therefore, although our original definition of academic preparedness could, in theory, bias our results, we have little reason to believe it does so in the empirical context of this article.

Policy Implications

By providing an empirical analysis of Black–White AP math disparities that is less vulnerable to common methodological issues of previous research, our study provides valuable policy implications. Given the evidence of preparedness-adjusted Black–White AP math enrollment disparities presented here—disparities which are not explained by Black–White differences in academic qualification for AP math courses—it follows that policy efforts to equalize opportunities for advanced enrollment in high school cannot be limited to diminishing racial disparities in early years of schooling. Although addressing the emergence of inequalities in early years is essential, the role of the enrollment process in producing racial disparities should not be ignored. We recognize that disparities that arise due to differences in academic preparation are not the only reason why policymakers might rethink how enrollment decisions are made and that these decisions could accommodate some level of affirmative action for historically disadvantaged groups as a way to mitigate a cycle of social inequalities. The fact that racialized feelings of social belonging in advanced course-taking environments tend to emerge when Black students are underrepresented in advanced courses is an important argument supporting this perspective (O’Connor et al. 2011; Tyson 2011). That said, we hope that the findings presented in this article may inform policy debates around racial inequalities in student coursework placement, and that the statistical approach we applied may prove useful for quantifying the extent to which selection processes tend to favor particular groups in other educational contexts.

Future Research

Constraints on the scope of our analyses also suggest directions for future work. First, the analysis in this article was restricted to AP math courses in one city, but estimating preparedness-adjusted placement disparities in different disciplines, different kinds of advanced courses, and different school districts may shed additional light on the extent to which enrollment processes favor some racial groups and can inform local policies and interventions. We believe that such investigations are particularly relevant given that, as demonstrated here, the traditional approach used to address these questions faces important methodological limitations and, thus, existing evidence runs the risk of misinforming policy and practice.

Second, our approach computes an overall measure of preparedness-adjusted enrollment disparities across races without differentiating between the multiple mechanisms shaping this quantity. Such mechanisms include the disparate impacts of formal and informal eligibility criteria for advanced courses, biased perceptions of students’ academic abilities, racial differences in parental involvement in placement decisions, and differences in students’ feelings of belonging in advanced course-taking environments. Distinguishing between these mechanisms is important for crafting effective policy interventions. Overall, we hope that the statistical approach we describe in this article may prove useful to account for the role of academic preparedness when measuring disparities in students' enrollment outcomes, and that our empirical findings can inform policy debates around racial inequalities in advanced coursework enrollment processes.

Funding Information

The authors acknowledge that they received no funding in support for this research.

Data and Code Availability Statement

The administrative data set used in this article was provided to the authors under a Data Use Agreement with NYC Public Schools (NYCPS) through the Research Alliance for New York City Schools Researchers may request access to these data from NYCPS at https://infohub.nyced.org/in-our-schools/working-with-the-doe/research-irb/faqs-for-external-data-requests. The code to reproduce all results in the paper is publicly available at https://github.com/joaosoutomaior/diff-acad-prep-AP-code. The code is organized so that anyone with access to the restricted administrative data could reproduce our results.

Notes

1 Some studies also adjust for selected sociodemographic variables in an attempt to distinguish the extent to which decisions favor students based on their race from the extent to which decisions favor students based on, for example, gender and/or social class.

2 We emphasize that we are not attempting to measure the "causal effect" of race on enrollment decisions, a concept that has been the subject of much academic debate (Gaebler et al. 2022; Greiner and Rubin 2011).

3 We note that the AP exam grading process, which is administered by the Education Testing Service (ETS), can substantially reduce the potential for readers' subjectivity (College Board 2023b; Education Testing Agency 2023). First, multiple-choice questions (which account for about half to two-thirds of AP scores) are graded by computers. Second, free-response questions are graded each June by hundreds of trained ETS readers (college faculty and high school teachers selected each year by ETS). Grading occurs during a seven-day period in selected onsite locations across the United States as well as remotely (since the Covid-19 pandemic). Completion of substantial training sessions is mandatory before the start of the reading period. We acknowledge, however, that subjectivity in grading might not be completely absent; for instance, the inclusion of names in students' free-response booklets could, in theory, reveal identity-related information.

4 To illustrate this general rhetoric, note that (Oakes and Guiton 1995) observe that: "(...) educators repeatedly expressed the wish to provide all students with courses in which they could be successful and maximize their potential. This was most evident when they talked about providing academic courses where low-ability students would not fail or feel pressure to drop out of school" (Oakes and Guiton 1995: 12).

5 We justify this assumption in more detail in Online Supplement B, where we provide a detailed description of our sample. Here, we simply note that it is rare for students in our data to take an AP math exam without taking at least one AP math course (just 2 percent of 12th grade students who do not enroll in any AP math courses take at least one AP math exam), or to take all their AP math courses before grade 11 (we removed the eight such students from our sample).

6 The emphasis on student enrollment in (and passage of) *at least one* AP math course (exam) is informed by research suggesting that benefits associated with AP coursework exist even for students who take (pass) only one AP course (exam) (Morgan and Klaric 2007; Fischer 2007; Chajewski et al. 2011).

7 Because taking the exam is necessary for passing the exam, $r_i(a,t) \equiv 0$ if $t = 0$; in contrast, a student could, in theory, pass an exam without enrolling in the advanced course, so $r_i(0,1)$ is not necessarily identically 0. However, for our analysis $r(1,1)$ is the sole potential outcome of interest.

8 In more detail, we use the implementation of XGBoost in the *xgboost* package in R (Chen et al. 2022). We implement a grid search over several combinations of parameters *max-depth, max-delt-step, min-child-weight, eta, and gamma* and use five-fold cross-validation to choose the parameters which maximize AUC. We impute missing values separately for the 2011 and 2012 cohorts according to variable means; however, no covariate is missing more than 2 percent of its values.

9 In Online Supplement E, we present a version of this figure where we disaggregate these distributions by students' AP math course enrollment and AP exam participation status (see Figure S.E.3).

10 Specifically, we generate 100 bootstrapped estimates of the race coefficients in Eq. 4, then add and subtract 1.96 times the estimated standard error of the bootstrapped estimates to our point estimate to obtain approximate 95 percent confidence intervals.

11 Balancing computational efficiency with search precision, we search over all combinations of the following parameters: $\alpha, \delta \in \{-\log(3), -\log(2), 0, \log(2), \log(3)\}$, and $q_{c_0}, q_{c_1} \in \{0, 0.1, \ldots, 0.9, 1\}$, where q_{c_0} denotes the prevalence of u among White students, and q_{c_1} denotes the prevalence of u among non-White students.

References

Ackerman, Phillip, Ruth Kanfer, and Charles Calderwood. 2013. "High School Advanced Placement and Student Performance in College: STEM Majors, non-STEM Majors, and Gender Differences." *Teachers College Record* 115:1–43.

Ainsworth, James W. 2002. "Why does it take a village? The mediation of neighborhood effects on educational achievement." *Social forces* 81:117–152.

Archbald, Doug and Elizabeth N. Farley-Ripple. 2012. "Predictors of placement in lower level versus higher level high school mathematics." *The High School Journal* pp. 33–51.

Ayres, Ian. 2005. "Three tests for measuring unjustified disparate impacts in organ transplantation: The problem of 'included variable' bias." *Perspectives in biology and medicine* 48:68–S87.

Benson, James and Geoffrey Borman. 2010. "Family, neighborhood, and school settings across seasons: When do socioeconomic context and racial composition matter for the reading achievement growth of young children." *Teachers College Record* 112:1338–1390.

Carter, Prudence L. 2005. *Keepin'it real: School success beyond Black and White*. Oxford, England: Oxford University Press.

Chajewski, Michael, Krista D. Mattern, and Emily J. Shaw. 2011. "Examining the role of advanced placement exam participation in 4-year college enrollment." *Educational Measurement: Issues and Practice* 30:16–27.

Champion, Joe and Vilma Mesa. 2016. "Factors affecting calculus completion among US high school students." *The Role of Calculus in the Transition from High School to College Mathematics, Washington, DC* .

Chen, Tianqi and Carlos Guestrin. 2016. "Xgboost: A Scalable Tree Boosting System." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, et al. 2022. "xgboost: Extreme Gradient Boosting." *R package version 1.5.2.1* .

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and economic opportunity in the United States: An intergenerational perspective." *The Quarterly Journal of Economics* 135:711–783.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2009. "The academic achievement gap in grades 3 to 8." *The Review of Economics and Statistics* 91:398–419.

College Board. 2023a. "Choosing Your AP Courses." Retrieved December 5, 2023 (https://apstudents.collegeboard.org/choosing-courses).

College Board. 2023b. "How are AP Exams scored?" Retrieved December 5, 2023 (https://apstudents.collegeboard.org/help-center/how-are-ap-exams-scored).

Condron, Dennis J. 2009. "Social class, school and non-school environments, and black/white inequalities in children's learning." *American Sociological Review* 74:685–708.

Conger, Dylan, Mark C. Long, and Patrice Iatarola. 2009. "Explaining race, poverty, and gender disparities in advanced course-taking." *Journal of Policy Analysis and Management* 28:555–576.

Domina, Thurston and Joshua Saldana. 2012. "Does raising the bar level the playing field? Mathematics curricular intensification and inequality in American high schools, 1982–2004." *American Educational Research Journal* 49:685–708.

Douglas, Daniel and Paul Attewell. 2017. "School mathematics as gatekeeper." *The Sociological Quarterly* 58:648–669.

Education Testing Agency. 2023. "AP Scoring Frequently Asked Questions." Retrieved December 5, 2023 (https://www.ets.org/careers/scoring-opportunities/ap-scoring/faq.html).

Engberg, Mark E. and Gregory C. Wolniak. 2010. "Examining the effects of high school contexts on postsecondary enrollment." *Research in Higher Education* 51:132–153.

Evans, Brent J. 2019. "How college students use advanced placement credit." *American Educational Research Journal* 56:925–954.

Fischer, E Mary J. 2007. "Settling into campus life: Differences by race/ethnicity in college involvement and outcomes." *The Journal of Higher Education* 78:125–161.

Fitzpatrick, Brian R. and Sarah Mustillo. 2020. "The right fit? Classroom mismatch in middle school and its inconsistent effect on student learning." *Sociology of Education* 93:277–294.

Fryer, Roland G. and Steven D. Levitt. 2006. "The black-white test score gap through third grade." *American Law and Economics Review* 8:249–281.

Gaebler, Johann, William Cai, Guillaume Basse, Ravi Shroff, Sharad Goel, and Jennifer Hill. 2022. "A causal framework for observational studies of discrimination." *Statistics and Public Policy* pp. 1–61.

Gamoran, Adam. 1992. "Access to excellence: Assignment to honors English classes in the transition from middle to high school." *Educational Evaluation and Policy Analysis* 14:185–204.

Gamoran, Adam and Robert D. Mare. 1989. "Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality?" *American journal of Sociology* 94:1146–1183.

Goldsmith, Pat Rubio. 2011. "Coleman revisited: School Segregation, Peers, and Frog Ponds." *American Educational Research Journal* 48:508–535.

Greiner, D. James and Donald B. Rubin. 2011. "Causal effects of perceived immutable characteristics." *Review of Economics and Statistics* 93:775–785.

Hacsi, Timothy A. 2004. "Document-based question: What is the historical significance of the Advanced Placement test?" *The Journal of American History* 90:1392–1400.

Hallett, Ronald E. and Kristan M. Venegas. 2011. "Is increased access enough? Advanced placement courses, quality, and success in low-income urban schools." *Journal for the Education of the Gifted* 34:468–487.

Hallinan, Maureen T. 1992. "The organization of students for instruction in the middle school." *Sociology of Education* pp. 114–127.

Hallinan, Maureen T. 1994. "Tracking: From theory to practice." *Sociology of education* 67:79–84.

Iatarola, Patrice, Dylan Conger, and Mark C. Long. 2011. "Determinants of high schools' advanced course offerings." *Educational Evaluation and Policy Analysis* 33:340–359.

Imbens, Guido W. and Donald B. Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Irizarry, Yasmiyn. 2021. "On Track or Derailed? Race, Advanced Math, and the Transition to High School." *Socius:: Sociological Research for a Dynamic World* 7:1–21.

Judson, Eugene, Nicole L. Bowers, and Kristi Glassmeyer. 2019. "Recruiting and encouraging students to complete advanced placement science and math courses and exams: Policies and practices." *Journal for the Education of the Gifted* 42:243–265.

Jung, Jongbin, Sam Corbett-Davies, Ravi Shroff, and Sharad Goel. 2019. "Omitted and included variable bias in tests for disparate impact." *arXiv preprint arXiv:1809.05651* .

Kelly, Sean. 2007. "The contours of tracking in North Carolina." *The High School Journal* 90:15–31.

Kelly, Sean. 2009. "The black-white gap in mathematics course taking." *Sociology of Education* 82:47–69.

Kelly, Sean and Heather Price. 2011. "The correlates of tracking policy: Opportunity hoarding, status competition, or a technical-functional explanation?" *American Educational Research Journal* 48:560–585.

Klugman, Joshua. 2013. "The Advanced Placement arms race and the reproduction of educational inequality." *Teachers College Record* 115:1–34.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2012. "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom." *The Economic Journal* 122:208–237.

Lee, Valerie E. and Anthony S. Bryk. 1988. "Curriculum tracking as mediating the social distribution of high school achievement." *Sociology of Education* pp. 78–94.

Lewis, Amanda E. and John B. Diamond. 2015. *Despite the best intentions: How racial inequality thrives in good schools*. Oxford University Press.

Lubienski, Sarah Theule. 2002. "A closer look at Black-White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data." *Journal of Negro Education* pp. 269–287.

Lucas, Samuel R. and Mark Berends. 2007. "Race and track location in US public schools." *Research in Social Stratification and Mobility* 25:169–187.

Lucas, Samuel R. and Adam Gamoran. 2002. "Track assignment and the black-white test score gap: Divergent and convergent evidence from 1980 and 1990 sophomores." *Closing the gap: Promising strategies for reducing the achievement gap* pp. 171–198.

Lucas, Samuel R., Santiago J. Molina, and John M. Towey. 2020. "Race/Ethnicity over Fifty Years of Structural Differentiation in K–12 Schooling: Period-Specific and Life-Course Perspectives." *Annual Review of Sociology* 46:355–378.

Malkus, Nat. 2016. "AP at scale: Public school students in Advanced Placement, 1990-2013." Technical report, The American Enterprise Institute.

Mickelson, Roslyn Arlin. 2001. "Subverting Swann: First-and second-generation segregation in the Charlotte-Mecklenburg schools." *American Educational Research Journal* 38:215–252.

Morgan, Rick and John Klaric. 2007. "AP Students in College: An Analysis of Five-Year Academic Careers. Research Report No. 2007-4." *College Board* .

Morton, Karisma and Catherine Riegle-Crumb. 2019. "Who Gets In? Examining Inequality in Eighth-Grade Algebra." *Journal for Research in Mathematics Education* 50:529–554.

NSB (National Science Board). 2020. "Vision 2030." Technical report, Report#: NSB-2020-15. [the easiest access to this source is via the URL].

Oakes, Jeannie. 2005. *Keeping track*. New Haven, CT: Yale University Press.

Oakes, Jeannie and Gretchen Guiton. 1995. "Matchmaking: The dynamics of high school tracking decisions." *American Educational Research Journal* 32:3–33.

O'Connor, Carla, Jennifer Mueller, R. L'Heureux Lewis, Deborah Rivas-Drake, and Seneca Rosenberg. 2011. "'Being' Black and strategizing for excellence in a racially stratified academic hierarchy." *American Educational Research Journal* 48:1232–1257.

Riegle-Crumb, Catherine and Eric Grodsky. 2010. "Racial-ethnic differences at the intersection of math course-taking and achievement." *Sociology of Education* 83:248–270.

Riegle-Crumb, Catherine, Sarah Blanchard Kyte, and Karisma Morton. 2018. "Gender and racial/ethnic differences in educational outcomes: Examining patterns, explanations, and new directions for research." In *Handbook of the sociology of education in the 21st century*, pp. 131–152. Springer.

Roscigno, Vincent J., Donald Tomaskovic-Devey, and Martha Crowley. 2006. "Education and the inequalities of place." *Social Forces* 84:2121–2145.

Rose, Heather and Julian R. Betts. 2004. "The effect of high school courses on earnings." *Review of Economics and Statistics* 86:497–513.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome." *Journal of the Royal Statistical Society: Series B (Methodological)* 45:212–218.

Schiller, Kathryn S. and Donald J. Hunt. 2011. "Secondary mathematics course trajectories: Understanding accumulated disadvantages in mathematics in grades 9–12." *Journal of School Leadership* 21:87–118.

Tyson, Karolyn. 2011. *Integration interrupted: Tracking, Black students, and acting White after Brown*. New York, NY: Oxford University Press.

Warne, Russell T., Gerhard Sonnert, and Philip M. Sadler. 2019. "The relationship between advanced placement mathematics courses and students' STEM career interest." *Educational Researcher* 48:101–111.

Xu, Di, Sabrina Solanki, and John Fink. 2021. "College acceleration for all? Mapping racial gaps in Advanced Placement and dual enrollment participation." *American Educational Research Journal* pp. 954–992.

Acknowledgements: We thank Johann Gaebler, Sharad Goel, Jongbin Jung and L'Heureux Lewis-McCoy for helpful comments and feedback. This study was conducted with data obtained through the Research Alliance for New York City Schools. The findings and conclusions are those of the authors and do not necessarily reflect the views of the Research Alliance.

João M. Souto-Maior: The Institute of Human Development and Social Change, New York University. E-mail: jms1738@nyu.edu.

Ravi Shroff: Department of Applied Statistics, Social Science, and Humanities, New York University. E-mail: ravi.shroff@nyu.edu.