Advanced Placement Gatekeeping and Racialized Tracking

Noah Hirschl[1] (iD) and Christian Michael Smith[2] (iD)

Abstract

Racialized tracking is central to sociological explanations for racially stratified educational outcomes. However, school officials' decision-making is of debated importance for explaining racialized tracking. We contribute to this literature by examining the effects of schools' enrollment policies for Advanced Placement (AP) courses. Using a unique combination of school survey data and administrative data from Wisconsin, we examine what happens to racial inequality in AP participation when school officials enforce performance-based selection criteria, which we call "course gatekeeping." We find that course gatekeeping has racially disproportionate effects. Although racialized differences in prior achievement partially explain the especially large negative effects among students of color, course gatekeeping produces Black-white and Hispanic-white disparities in participation even among students with similar, relatively low prior achievement. We further find that course gatekeeping has longer-run effects, particularly discouraging Black and Asian or Pacific Islander students from attending highly selective four-year colleges.

A salient feature of the U.S. secondary school system has been de facto racial segregation across curricular levels within schools, a phenomenon commonly referred to as racialized tracking (Clotfelter, Ladd, and Vigdor 2003; Mickelson 2001; Tyson 2011). Although students are no longer sorted into a strictly linked set of courses through formal tracks, course enrollment remains highly racially segregated (Lucas 1999; Lucas and Berends 2007). Scholars of education and race have identified racialized tracking as one of the key mechanisms that contributes to the ongoing production of racial inequality in educational outcomes (Conwell 2020; Diamond 2006).

We join other scholars by investigating the mechanisms that lead to racialized tracking and their consequences for students (Lewis and Diamond 2015; Lewis-McCoy 2014; Tyson 2011). In particular, we test whether high schools

exacerbate racial inequality when they impose performance-based eligibility criteria for advanced coursework. Existing literature points to different predictions regarding whether and why these eligibility criteria exacerbate racial inequality. With a longitudinal research design that capitalizes on a combination of administrative and survey data from the state of Wisconsin, we weigh the evidence supporting each prediction.

We focus on Advanced Placement (AP) courses, a curricular program that has come to

[1]University of Wisconsin-Madison, Madison, WI, USA
[2]University of California, Merced, CA, USA

Corresponding Author:
Christian Michael Smith, University of California, Merced, 5200 Lake Road, Merced, CA 95343, USA.
Email: csmith97@ucmerced.edu

dominate the upper strata of coursework in public high schools in the United States. The AP program is a set of more than 30 subject exams that the College Board constructs to reflect the rigor of college-level courses. Nearly 3 million high school students sat for over 5 million AP exams in 2019, a remarkable expansion since the program's inception as a curriculum for students bound for the most elite colleges (College Board 2019; Schneider 2009). AP aims to give students experience with college-level coursework, and if they attain a score of at least a 3 out of 5 on the standardized end-of-year exam, they may earn college credit, often allowing them to skip certain classes, expand their coursework breadth and depth, and graduate in less time (Avery et al. 2018; Evans 2019; Gurantz 2019; Smith, Hurwitz, and Avery 2017). Although opportunities to take AP courses have become more widespread across schools, the AP program has largely replicated earlier patterns of curricular segregation by race within schools (Kolluri 2018; Malkus 2016; Price 2021).

We study what happens to racial inequality in AP participation when schools change their policies regarding which students are permitted to participate. We also study what happens to racial inequality in college attendance and selective college attendance when schools make these changes; these postsecondary outcomes may be downstream consequences of AP policies because the rigor of a student's high school curriculum plays a key role in college admissions, especially at the most selective colleges (Bastedo, Howard, and Flaster 2016; Bowman and Bastedo 2018). We use a unique linkage between school survey data and administrative data from Wisconsin as our empirical case. By focusing on the effects of changes in schools' policies over time, our study circumvents many forms of statistical bias that other observational studies cannot.

In addition to advancing theory on the links between race, education, and stratification, our study informs policy and practice meant to reduce racialized tracking in high schools. This practical contribution is important because racial inequality in AP participation may have larger implications. First, AP courses may foster enrollment in baccalaureate colleges generally and selective colleges particularly, which may be important for the educational trajectories and later life outcomes of racially minoritized students (Cortes 2010; Melguizo 2010; Small and Winship 2007). Second, we emphasize, along with other scholars, that

racialized AP tracking may have other, harder to measure costs. Many argue that racialized tracking is a form of organizational racism that materializes racial hierarchy in the everyday structure of schools, in turn heightening students' perception of the link between racial identity and academic success (Lewis and Diamond 2015; Ray 2019; Tyson 2011) and subjecting some Black students to accusations that they ''act white'' (O'Connor et al. 2011).

BACKGROUND

Organizational decisions about who is in class with whom reflect deeper tensions between schools' ability to exacerbate or alleviate social inequalities. The consensus view of tracking and inequality can be summarized simply: Stronger tracking systems tend to increase the variance in learning, thus potentially reinforcing racial inequality in educational outcomes (Gamoran and Mare 1989; Hout and DiPrete 2006). Tracking practices have been used as a method of racial exclusion in educational institutions in the United States. For example, racialized tracking systems were one way that schools reinstituted racial segregation in response to court-mandated school integration (Tyson 2011). Thus, tracking practices have been a focus of scholars and policymakers seeking to identify and reduce racial inequality in educational outcomes.

Contemporary school tracking systems vary widely. Sørensen (1970) defined a set of theoretical dimensions that characterize diversity across tracking systems, and his theoretical model has been further expanded in later empirical applications (Domina et al. 2019; Gamoran 1992; Kelly 2007; Rosenbaum 1976). In this article, we define and study the effects of a particular aspect of high schools' tracking systems on racial inequality in course-taking: school officials' use of performance-based selection criteria to control enrollment, which we call ''course gatekeeping,'' or ''gatekeeping'' for short.[1]

In practice, schools use a variety of mechanisms that fall under this definition of course gatekeeping. Typical gatekeeping measures include ''objective'' enrollment benchmarks, such as minimum grades or test scores, and ''subjective'' enrollment criteria, such as teacher recommendations or vague behavioral requirements (Kelly 2007; Kelly and Price 2011). Prerequisite
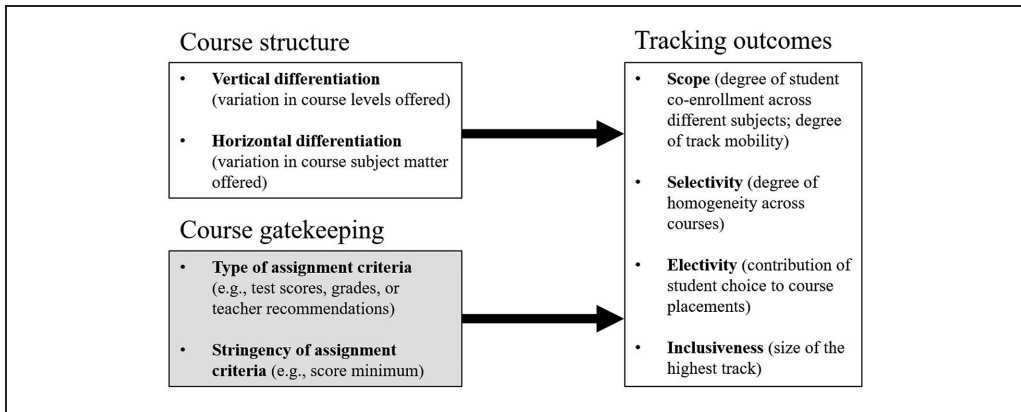
Figure 1. Conceptual model of the relationship between course gatekeeping and tracking outcomes.

or corequisite courses are also common entry criteria for upper-level coursework. In the case of these policies, we draw a distinction between requirements that are intended to ensure minimum prior content knowledge (e.g., requiring that students have taken two years of high school English before enrolling in AP English), which we would not call course gatekeeping, and those requiring a particular level of a prior course that covers similar content to another course (e.g., requiring that students take ''honors'' instead of ''standard'' English before enrolling in AP English), which we would call gatekeeping because those policies more clearly use performance-based criteria. Throughout this article, we term the absence of any gatekeeping measures ''open enrollment.''

In Figure 1, we build on Sørensen's and others' dimensions of organizational differentiation to locate course gatekeeping in a conceptual model of tracking outcomes. Course gatekeeping depends on what Sørensen (1970) calls the assignment mechanism: the types of assignment criteria and their stringency. Teachers may gatekeep or not, and conditional on gatekeeping, they may do so to varying degrees depending on the exclusivity of the criterion (e.g., by specifying a specific minimum grade in a prior class).

The conceptual model in Figure 1 also clarifies the object of our inquiry by distinguishing between the policies and practices that schools use to produce differentiation (''course gatekeeping'' and ''course structure'' on the left side of Figure 1), and the actual degree of differentiation (''tracking outcomes'' on the right side of Figure 1). Course gatekeeping and structure are the two primary policy means by which schools affect organizational differentiation. This distinction avoids conflating the effects of policies with other factors producing tracking outcomes, such as the size and composition of the student body, students' own enrollment choices, or scheduling conflicts (see Domina et al. 2019; Kelly 2007). Thus, course structure and course gatekeeping are logically prior to tracking outcomes and have imperfect relationships with tracking outcomes.

In making this conceptual distinction between policies and outcomes, we cleanly separate gatekeeping from two closely related, canonical dimensions of tracking systems: electivity and selectivity. Sørensen (1970:361, emphasis added) originally defined electivity as ''the degree to which students' own decisions *are allowed* to be a determining factor in the assignment to groups,'' which seems to define the concept as a school policy or practice. In contrast, we define electivity as it is typically used in subsequent empirical work: the realized ''contribution'' of student choice to course placements given the set of constraints imposed by school policies (Kelly 2007:18; see also Gamoran 1992). We generally expect course gatekeeping to limit electivity. However, course gatekeeping may have little effect on electivity if students strictly self-sort into courses according to their prior performance. In that case, student enrollment patterns would be mostly unaffected by a performance-based barrier.

Similarly, Sørensen's (1970:362) original definition of selectivity is "the amount of homogeneity that school authorities intend," which again evokes school policies instead of outcomes. However, most empirical work operationalizes selectivity as the *observed* degree of homogeneity across course or track levels (e.g., Domina et al. 2019; Gamoran 1992). We follow these authors to define selectivity as an observed outcome rather than an intended outcome, or policy. Thus—as with electivity—course gatekeeping will only affect selectivity to the extent that it changes students' enrollment behavior.

Course Gatekeeping and Racialized Tracking

How and to what extent should we expect course gatekeeping to affect racialized tracking? Empirically, there are three possibilities: Gatekeeping may have very little or no effect, it may exacerbate racialized tracking, or it may ameliorate racialized tracking. Our main objective in this article is to provide novel evidence to assess which of these predictions sees support.

First prediction: Minimal effect of gatekeeping.
To support the prediction that school officials' gatekeeping has little or no effect, scholars have suggested that other racialized, interlocking forces keep students racially segregated. In this scenario, even if teachers remove or enact a stringent requirement for access to upper-level courses, it has no effect on students' enrollment because students tend to self-segregate anyway (Yonezawa, Wells, and Serna 2002). Prior work suggests these interlocking forces include students' academic self-concept, social networks, informational barriers, hostile experiences in advanced courses, and parental lobbying efforts.

As a strategy to "protect their sense of competence," many students choose classes in which they expect to succeed (Tyson 2011:139). By the time they reach high school, students' sense of their academic competence has inflated or deflated according to accumulated, racialized academic experiences, such as earlier placement in gifted and talented programs; reinforcement from parents, teachers, and peers; and a long track record of academic success (Tyson 2011; Yonezawa et al. 2002). Thus, students may have already restricted their course choices regardless of how teachers admit students into advanced classes.

Furthermore, students tend to prefer taking classes with their friends, and friendship networks in high schools are typically highly racially segregated (Moody 2001). Black and other racially minoritized students may be reluctant to enroll in classes where few of their friends are present (Francis and Darity 2020).

Relatedly, when very few Black or Hispanic students are enrolled in a given advanced class, such as an AP course, those few report feeling uncomfortably visible in a white-dominated space and sensing that their peers devalue their perspective (Lewis and Diamond 2015; Taliaferro and DeCuir-Gunby 2008; Tyson 2011; Yonezawa et al. 2002). Similarly, O'Connor et al. (2011: 1232) argue that a "racially stratified academic hierarchy" in a high school makes high-achieving Black students more vulnerable to accusations of "acting white." Thus, one can imagine AP courses being undesirable to racially minoritized students in such schools regardless of the gatekeeping structures in place.

Moreover, some parents may help make gatekeeping policies moot by working around whatever policies are in place. White parents—especially those of high socioeconomic status—are apt to challenge school policies (Lewis-McCoy 2014), and schools are apt to capitulate, including when it comes to course-placement challenges (Calarco 2020). Even if these parents do not challenge the school policies, they may work hard to adapt to the policies (Alon 2009; e.g., by investing in resources that help their children reach the eligibility criteria).

A study by Yonezawa et al. (2002) is perhaps most germane to the prediction that gatekeeping will have minimal effects. They argue that a combination of the aforementioned factors prevented any change in racialized tracking when the set of middle and high schools in their qualitative study allowed students complete freedom to choose their courses.

Second prediction: Gatekeeping exacerbates racialized tracking.
The existing evidence for this prediction suggests three potential mechanisms. These mechanisms are not, in general, mutually exclusive, and components of each are likely relevant in different contexts.

First, a set of quantitative studies suggests that gatekeeping in advanced courses excludes

minoritized students chiefly because they exclude students with lower prior academic achievement. These studies find that prior academic achievement accounts for most, if not all, of the racial differences in high school course-taking, particularly in the past few decades (Conger, Long, and Iatarola 2009; Gamoran and Mare 1989; Kelly 2009; Lucas, Molina, and Towey 2020). This finding may not hold in all contexts: Schools with a high degree of socioeconomic or racial/ethnic heterogeneity tend to have more pronounced inequalities net of achievement (Lucas and Berends 2002, 2007; Oakes 1994). But the broader finding that earlier achievement disparities can account for racial disparities in advanced course placement suggests that schools strictly assign students to courses based on their prior achievement, regardless of racial identity. This further suggests that tightening entry criteria for advanced courses will affect racialized tracking if some racialized groups are more concentrated in the academic strata that are excluded under the new criteria.

Second, gatekeeping will reinforce racialized tracking if teachers and counselors hold systematically lower academic expectations for Black and Hispanic students, leading to default enrollment in lower-track classes, racial bias in faculty recommendations for AP courses, or higher resistance to students' or parents' attempts to appeal course placements (Lewis and Diamond 2015; Yonezawa et al. 2002). Some studies support the potential role of racialized teacher expectations, particularly with respect to Black students. Evidence from studies of twentieth-century teachers shows they expected worse academic performance from Black students than they expected from white students with identical records (for a review, see Ferguson 2003). More recently, Grissom and Redding (2016) found that elementary school teachers are less likely to refer Black students to gifted programs than they are to refer observationally equivalent white students. In high schools, teachers' educational expectations for Black students are dramatically lower when the teacher is non-Black rather than Black (Fox 2016; Gershenson, Holt, and Papageorge 2016). This bodes poorly for Black students given that the majority of teachers are non-Black (Schaeffer 2021), even at schools with majority-Black student bodies (Spiegelman 2020). Fox (2016), Gershenson and colleagues (2016), and Grissom and Redding (2016) all

highlight the importance of student-teacher racial match for Black students, but none of the studies find similar patterns for students in other racial groups.

Similarly, Francis, De Oliveira, and Dimmitt (2019) found that compared to fictional students with transcripts that had no name or a name coded demographically differently, fictional students with an otherwise identical transcript that had a Black female-coded name were far less likely to receive school counselors' recommendation to take AP calculus. However, Francis and colleagues (2019) did not find that Black male-coded transcripts were similarly disadvantaged. In a study investigating teachers' recommendations for AP courses, Fox (2016) found that student-teacher racial match did not matter for the likelihood of AP recommendation, even though it mattered significantly for educational expectations (among Black students but not among students in other racial groups). However, Fox's (2016) study concerns the effect of student-teacher racial match rather than the effect of student race per se, and thus it does not directly indicate whether teachers are more inclined to recommend AP courses to white students than comparable Black students. In summary, the existing literature shows that school counselors tend to discriminate against Black female students in AP course recommendations, and it gives some suggestive clues that teachers might expect less achievement and attainment from Black students than from comparable white students. However, the field lacks clear evidence on whether AP course recommendations are applied racially unequally on a broad level.

Third, gatekeeping might affect racialized tracking due to racial differences in students' propensity to choose to enroll in AP courses among relatively low-achieving students. Gatekeeping only excludes students who both want to enroll in AP courses and whose achievement falls below the assignment criterion's threshold. Thus, if any racial group is overrepresented in the subpopulation that meets those two conditions, the effects of gatekeeping will be more severe for that group. In particular, Black students (and perhaps other nonwhite racialized groups) express more pro-school attitudes and are more likely to aspire to and attend baccalaureate colleges than are white students with similar achievement levels (Bennett

and Xie 2003; Charles, Roscigno, and Torres 2007; Downey, Ainsworth, and Qian 2009; Mickelson 1990). If these phenomena lead relatively low-achieving minoritized students to be more inclined to enroll in AP courses than relatively low-achieving white students, on average, then gatekeeping may have racially disproportionate effects.[2]

Third prediction: Gatekeeping ameliorates racialized tracking. Although relatively little evidence supports this prediction, it is possible that a new or more stringent assignment mechanism could promote racial equality in AP enrollment. In one empirical example of a similar phenomenon, Card and Giuliano (2016) found that when universal screening for gifted education replaced a system based on parents' and teachers' referrals, Black and Hispanic students' representation in the program increased. This case suggests that when an existing system of allocation is very strongly biased against minoritized students, new performance-based criteria can reduce inequality.

Postsecondary Implications of Course Gatekeeping

Why might AP gatekeeping policies affect enrollment in baccalaureate colleges generally and selective colleges particularly? First, enrollment in AP courses may affect college plans, giving college-aspiring students a sense they can achieve in a rigorous college context, raising their expectations for the kind of institution they can attend (Karlson 2015). On the other hand, if students struggle academically in AP courses, it could erode their confidence in their ability to complete difficult coursework, thus reducing college enrollment (Conger et al. 2021). Second, AP courses improve students' chances of admission to highly selective colleges (Bastedo et al. 2016). Almost half of admissions officers at selective colleges indicate that the rigor of high school coursework is the most important piece of information for determining applicants' academic merit, and at the very most selective colleges, over two-thirds of admissions officers say so (Bowman and Bastedo 2018). Therefore, expanding AP access could facilitate selective college admission particularly.

THE PRESENT STUDY

The main objective of the present study is to bring new evidence to clarify the role of gatekeeping in racialized tracking. We use data from public schools in Wisconsin to investigate the effects of changes in schools' AP enrollment policies on participation in those courses. We then examine whether these changes have downstream effects on college-going behavior. Our study adds to existing research by explicitly testing the strength of course gatekeeping as a mechanism of racialized tracking. Although other scholars have thoroughly described the sources and consequences of racialized tracking, our study is the first to estimate the effects of course gatekeeping specifically, and we do so using a school fixed-effects design and unique, high-quality data. It is also the first to study this mechanism in the context of the growing AP program and to follow students into college to examine downstream implications for postsecondary enrollment. Our contributions thus have relevance to sociological theory aiming to explain how schools produce racial inequality and to policy decisions that allocate access to advanced coursework.

DATA AND METHODS

Data

This study uses data from two sources: Wisconsin's Statewide Longitudinal Data System (SLDS) and the Office of Civil Rights Data Collection (OCRDC) surveys. The SLDS is a census of Wisconsin public school students that tracks these students across their educational careers. It includes information on demographics, enrollments, test scores, grades, and other educational outcomes. We study students who attended the 327 public traditional and charter high schools in Wisconsin that offered at least some AP courses and at least some students took AP exams for the duration of the study between the 2013–2014 and 2017–2018 school years. Our target population consists of three cohorts of first-time 10th graders[3] at these schools in 2013–2014, 2014–2015, and 2015–2016. These students had expected high school graduation dates in 2016, 2017, and 2018. Our analytic sample contains 144,669 students after the adjustments described in the remainder of this section.

Our first outcome of interest is students' participation in AP. We link our administrative data with College Board records to measure all the AP exams a student attempts. Our key outcome is simply a count of the total number of AP exams students took in Grades 10 to 12. Unfortunately, the available data do not allow us to directly measure AP course enrollment, which is a more proximate outcome of AP gatekeeping. Not all students who take a course take the exam: Nationally, about two-thirds to four-fifths of students who enroll in a given AP course ultimately take the exam (Malkus 2016; Price 2021). Under plausible assumptions, using exam counts as a proxy yields no bias in racial/ethnic interaction terms and no bias in, or even conversative estimation of, the group-specific gatekeeping effects we estimate. See Appendix B in the online supplement for a more thorough treatment of this issue.

Our secondary outcome of interest is students' college enrollment. We measure students' college enrollment in the fall after their senior year through a linkage between our administrative records and college enrollment from the National Student Clearinghouse.[4] We analyze students' enrollment in any four-year institution, and among four-year institutions, we also examine enrollment in "highly selective" schools, which we define as one of the 236 institutions that the *2009 Barron's Profiles of American Colleges* (College Division of Barron's Profiles of American Colleges 2008) ranked as "very competitive plus" or higher.

Our main treatment variable is schools' AP gatekeeping policies. To measure gatekeeping policies in this period, we merge school-level SLDS data with OCRDC surveys from the 2013–2014, 2015–2016, and 2017–2018 school years. This biennial survey is conducted by the federal Office of Civil Rights and has drawn responses from the universe of public education agencies in the country since 2012. Response rates are perfect because schools are required by law to fill out the survey.

In the latest three OCRDC surveys, schools were asked whether "a student is allowed to enroll in *all AP courses that the school offers* . . . without needing a recommendation or without meeting other criteria (except for any necessary course prerequisites)" [emphasis added]. We use this measure as a proxy for schools' AP gatekeeping policies. However, this survey response might fall short of measuring all aspects of gatekeeping as we have conceptualized it in three ways. First, the question asks whether schools engage in gatekeeping in any of their AP courses. For instance, if a single teacher enforces a prior grade requirement for their class, the school is supposed to answer "no" on the survey. Thus, this measure may capture a narrower range of gatekeeping policies than exists across schools. Second, it excludes prerequisites, which we consider in certain cases to qualify as course gatekeeping measures (see aforementioned description). However, the other requirements measured by this survey question, such as prior grades, test scores, and recommendations, are highly variable across schools: According to one 2002 survey of AP teachers, about 50 percent of teachers required a minimum grade, and nearly 60 percent required a recommendation (Milewski and Gillie 2002). Both in Wisconsin and in the national OCRDC survey data from 2017–2018, about 30 percent of all high schools reported using AP gatekeeping of some kind. Thus, we measure an important aspect of between- and within-school variation in AP gatekeeping practice. Third, our measure does not record the type or stringency of the gatekeeping measure in place, so we cannot separate the effects of different types of gatekeeping measures in this study.

Because the exact same question was included across survey years, we use changes in schools' responses to indicate a change in policy over time. We use the terms "AP gatekeeping" to refer to the policy implied by answering "no" to the survey question and "open enrollment" to refer to the policy implied by "yes." We constructed our measure of gatekeeping by averaging the enrollment policy that each student experienced (1 = gatekeeping, 0 = open enrollment) in Grades 10 to 12.[5] Thus, our measure varies across students from 0 (always open enrollment) to 1 (always gatekeeping). For example, if Student A attended Grades 10 to 12 in a school when the policy was always gatekeeping, the measure would equal 1 for Student A. If Student B attended Grades 10 to 12 at the same school three later years and the school allowed open enrollment in one of those three years, the measure would equal $0.\overline{66}$ for Student B.

There is enough variation in schools' policies over time to identify the effects of interest. Figure 2 displays trajectories of school policy change over the three waves of the OCRDC survey. Open enrollment policies are more common than gatekeeping policies across this entire period in Wisconsin. However, there is considerable flux
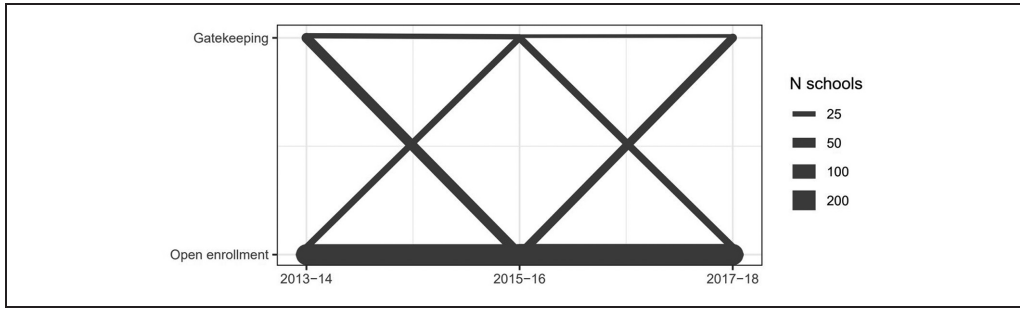
Figure 2. Change in high schools' AP enrollment policies by OCRDC survey school year.
Source: Data are drawn from the Wisconsin State Longitudinal Data Systems and the OCRDC.
Note: Each line segment's width is proportional to the number of schools with that sequence of AP enrollment polices between the school years. AP = Advanced Placement; OCRDC = Office of Civil Rights Data Collection.

over time. In the two gaps between OCRDC survey waves, 71 and 79 high schools changed their policy in either direction. This degree of variation is consistent with other work that shows considerable variation in tracking practices both across and within schools across time (Domina et al. 2019). Our research design exploits these changes in schools' policies.

Research Design and Estimation

A simple comparison of AP participation between schools with different policies could be subject to confounding by unobserved school-level differences because schools' course-placement policies differ in systematic ways (see Spade, Columba, and Vanfossen 1997). For instance, schools with open enrollment policies might also be more likely to provide other resources and programming intended to reduce disparities among racial groups, thus confounding naïve comparisons between schools. We mitigate this source of bias by measuring the effects of AP gatekeeping among cohorts of 10th graders exposed to different policies in the same school using school fixed-effects models.

We chose to estimate Poisson models because our outcome of interest is a strictly nonnegative count, so we expect the log-link functional form to be the most appropriate characterization of the relationship between the predictors and the conditional expectation of the outcome.[6] These models take the following general form:

$$\log AP_{ijk} = \alpha + \beta G_{ijk} + \sum_l \gamma_l X_{ijkl} + \theta_j + \tau_k + \varepsilon_{ijk},$$

where AP_{ijk} is the number of AP exams taken by student i in school j in cohort k. The measure G_{ijk} is the averaged gatekeeping policy for each student described in the previous section, which varies between 0 and 1. The parameter β captures the effect of experiencing AP gatekeeping in all years versus open enrollment in all years. In subsequent models, we include additional interaction terms that allow β to vary by race-ethnicity and by prior achievement to test the specific hypotheses described previously. In models with race interactions, we omit the racial category ''non-Hispanic white'' from the interaction set. The γ_l parameters capture the effects of a set of covariates X, which are indexed by l and measured for each student i. The θ_j are school fixed effects, and the τ_k are cohort fixed effects. We cluster the standard errors at the school level. In the final section of the analysis, we estimate parallel models for college attendance outcomes. These models are the same in all respects except that we estimate them as logit models for binary outcomes: any four-year college attendance and highly selective four-year college attendance.

Covariates

Our main interest is whether the effects of AP gatekeeping differ by race and ethnicity. Our administrative measure of race/ethnicity categorizes students as non-Hispanic white, non-Hispanic Black, Hispanic, non-Hispanic Asian or Pacific Islander, non-Hispanic American Indian or Alaska Native, or two or more races. We do not aggregate this measure further in our analysis.

The Hispanic and Asian populations in our sample are heterogeneous in terms of nativity and immigrant generation, with implications for the cultural, economic, and social capital students bring to school. In Wisconsin, three in four Hispanic individuals report Mexican descent, and most of the remainder report Puerto Rican descent. The predominant Asian immigrant group in Wisconsin is Hmong: About one in three individuals identifying as Asian are Hmong, 18 percent are Asian Indian, and 13 percent are Chinese (Curtis and Lessem 2014). These population characteristics bear on the social processes underlying our results and on the extent to which our results generalize to other contexts.

We use measures of students' prior achievement as further control variables and moderators. We construct our key measure of prior achievement using students' GPAs when they were in ninth grade. Our secondary measure of prior achievement is students' test scores on their eighth-grade Wisconsin Knowledge and Concepts Examinations in math and language arts, the most recent standardized tests that students in our sample took. We standardize these scores so that mean = 0 and standard deviation = 1. We expect that in addition to GPA and test performance, students' prior track location is an important determinant of whether they take AP courses and is likely a factor in gatekeeping decisions (Kelly and Carbonaro 2012). Unfortunately, our coursework data do not reliably identify course subjects or levels. We therefore construct a proxy for track location in ninth grade using the average math and English language arts test scores of students' ninth-grade classroom peers, which we expect strongly correlates with students' overall position in their schools' tracking system.[7] Test scores and GPA are the only sources of missingness in our data: 6 percent of students lack a ninth-grade GPA, 12.5 percent lack eighth-grade test scores, and 0.5 percent could not be matched to classmates with nonmissing test scores to calculate our measure of track location. In total, 15.8 percent of student records lack one of these key achievement measures, and we drop these students from our sample.[8]

In addition to the race/ethnicity and achievement measures described previously, all models include controls for binary sex, whether the student is an English-language learner, whether the student is classified as having a disability, and a measure of economic disadvantage. We measure

economic disadvantage using receipt of free or reduced-price lunch. Following evidence that longitudinal measures capture more meaningful variation (Michelmore and Dynarski 2017), we use a measure of the proportion of years we observe students receiving free or reduced-price lunch in the years they were enrolled in Wisconsin public schools.

We also control for changes over time in students' opportunity to take AP exams by measuring the number of distinct AP exam subjects that schools offered while students were enrolled in them. We take the average over students' enrollment in Grades 10 to 12 and log transform it. By controlling for course offerings, we neutralize the possibility that any expansions in course catalogs could mechanically increase AP participation for some groups and not others (Rodriguez and McGuire 2019). Finally, we control for school-by-cohort averages of students' test scores and prior GPA. These measures account for the possibility that schools or teachers gatekeep based on the composition of students in each cohort. For instance, schools might be more likely to use course gatekeeping measures to allocate seats when a particularly high-achieving cohort of students overenroll in AP courses.

RESULTS AND DISCUSSION

Racial Inequality in AP Participation

Patterns of AP exam participation reflect racialized tracking in Wisconsin's high schools. On average, white students take nearly one more AP exam than do Black or American Indian students and about one-half more exams than the average Hispanic student takes (see Table 1). On average, Asian/Pacific Islander students take more exams than white students, although the variability in the number taken is also highest in that group. Students from two or more racial groups fall near the middle in terms of the average number of exams taken.

We also see notable racial differences in the experience of gatekeeping policies. White, Asian/Pacific Islander, and American Indian students are most likely to experience gatekeeping, at 18 and 19 percent of their 10th- to 12th-grade years, on average, whereas Hispanic students and Black students experience gatekeeping 14 and 9 percent of the time, respectively. The

Table 1. Means (Standard Deviations) of Measures, by Race/Ethnicity.

	Racial/ethnic identification					
	White	Black	Hispanic	Asian/Pacific Islander	American Indian	Two+ races
Number of AP exams taken	1.23	0.38	0.62	1.85	0.34	1.00
	(2.04)	(1.10)	(1.47)	(2.94)	(1.05)	(2.02)
Took ≥ 1 AP exam	0.40	0.16	0.24	0.45	0.15	0.31
Number of AP exams passed	0.84	0.11	0.31	1.25	0.18	0.70
	(1.76)	(0.60)	(0.99)	(2.60)	(0.75)	(1.76)
Attended a four-year college	0.43	0.18	0.19	0.44	0.15	0.33
Attended a highly selective four-year college	0.10	0.02	0.04	0.18	0.03	0.11
AP subjects offered at school	13.7	13.6	14.7	16.4	11.7	15.6
	(6.5)	(6.5)	(5.9)	(5.8)	(5.7)	(6.3)
Average AP gatekeeping policy	0.18	0.09	0.14	0.19	0.19	0.15
	(0.30)	(0.22)	(0.28)	(0.32)	(0.30)	(0.28)
9th-grade GPA	2.97	2.01	2.35	3.15	2.20	2.64
	(0.88)	(0.95)	(0.97)	(0.79)	(0.96)	(1.02)
8th-grade WKCE math score (standardized)	0.17	−0.89	−0.46	0.09	−0.47	−0.08
	(0.91)	(1.03)	(0.94)	(1.06)	(0.98)	(1.01)
8th-grade WKCE ELA score (standardized)	0.12	−0.67	−0.40	−0.06	−0.41	−0.06
	(0.96)	(0.92)	(0.88)	(0.98)	(0.91)	(1.02)
Mean 9th-grade classmates' standardized math scores	0.17	−0.51	−0.22	0.11	−0.17	0.06
	(0.39)	(0.59)	(0.49)	(0.52)	(0.44)	(0.46)
Mean 9th-grade classmates' standardized ELA scores	0.08	−0.46	−0.24	0.03	−0.19	−0.01
	(0.34)	(0.48)	(0.42)	(0.46)	(0.38)	(0.41)
Proportion years FRL	0.24	0.78	0.70	0.52	0.63	0.48
	(0.35)	(0.33)	(0.37)	(0.43)	(0.39)	(0.43)
Female	0.49	0.49	0.49	0.51	0.50	0.50
Classified as disabled	0.11	0.22	0.13	0.07	0.21	0.16
English-language learner	0.00	0.01	0.17	0.13	0.01	0.01
Number of students	112,328	9,931	12,926	5,052	1,337	3,095

Source: Data are from the Wisconsin State Longitudinal Data Systems merged with survey data from the Office of Civil Rights Data Collection.
Note: AP = Advanced Placement; WKCE = Wisconsin Knowledge and Concepts Examinations; ELA = English language arts; FRL = receipt of free or reduced-price lunch.

Table 2. Characteristics of Schools with Differing AP Enrollment Policy Sequences.

	Changes policy	Always gatekeeping	Always open enrollment
Student characteristics			
White	0.81	0.76	0.76
Black	0.05	0.04	0.08
Hispanic	0.08	0.10	0.09
Asian/Pacific Islander	0.03	0.07	0.04
American Indian	0.01	0.01	0.01
Two+ races	0.02	0.02	0.02
Proportion years FRL	0.32	0.36	0.34
AP exams taken	1.06	0.75	1.18
School characteristics			
AP subjects offered	13.3	11.6	14.5
Annual enrollment	670	668	704
Number of students	52,244	4,936	87,489
Number of schools	122	11	194

Source: Data are from the Wisconsin State Longitudinal Data Systems merged with survey data from the Office of Civil Rights Data Collection.
Note: Cells contain proportions or means. AP = Advanced Placement; FRL = receipt of free or reduced-price lunch.

particularly low rate of gatekeeping experienced by Black students reflects the disproportionate enrollment of Black students in Milwaukee and Madison, both urban districts with open enrollment policies throughout the study period. Table 2 shows the characteristics of schools with differing gatekeeping policy sequences across the study period. If schools that change their policies are very different from other schools, it may limit the external validity of our findings.[9] We find that those schools are generally similar in terms of poverty rate, enrollment, and AP participation to other schools in Wisconsin, with the exception that ''always gatekeeping schools'' tend to offer fewer AP subjects and students take fewer exams.

Gatekeeping and Racial Inequality in AP Participation

Our findings suggest that AP gatekeeping policies make access to these advanced courses modestly more exclusive. The first column in Table 3 displays the main effect of gatekeeping from our specification of Model 1. The overall estimated effect of gatekeeping is to reduce AP exam-taking by 13 percent, on average $(1 - e^{-0.14} = 0.13)$. This point estimate is statistically significant at the 1 percent level.

However, the average estimated effect conceals important heterogeneity. In Model 2, we find some evidence of racially disproportionate effects of gatekeeping: These policies seem to especially reduce AP exam-taking among Hispanic and perhaps among Black students and other students of color. While gatekeeping reduces exam counts by an estimated, statistically significant 11 percent, on average, among white students $(1 - e^{-0.12} = 0.11)$, it reduces exam counts by 24 percent $(1 - e^{-0.12-0.16} = 0.24)$ among Black students and 28 percent $(1 - e^{-0.12-0.21} = 0.28)$ among Hispanic students. However, only the white-Hispanic effect difference is statistically significant at conventional levels; there is considerably more uncertainty in the white-Black contrast, in part due to the smaller number of Black students who experience gatekeeping regimes. The interaction terms are also somewhat negative among Asian or Pacific Islander students and multiracial students and slightly positive among American Indian/Alaska Native students; none are statistically distinguishable from zero.

Given the uncertainty in many of these interaction effects, the results in Model 2 of Table 3 provide some qualified support for our second prediction: Gatekeeping appears to exacerbate racialized tracking, particularly between Hispanic and white students. However, we identified multiple possible

Table 3. Estimated Effects of AP Gatekeeping Policy on AP Exams Taken from Fixed-effects Poisson Models, Heterogeneous Effects by Race/Ethnicity and Prior Achievement.

	Total number of AP exams taken			
	(1)	(2)	(3)	(4)
Main effect				
AP gatekeeping	−0.14** [−0.24, −0.03]	−0.12* [−0.22, −0.01]	−0.10+ [−0.20, 0.00]	−0.33** [−0.53, −0.13]
Race interactions				
AP gatekeeping × white		—		—
AP gatekeeping × Black		−0.16 [−0.49, 0.17]		−0.03 [−0.34, 0.29]
AP gatekeeping × Hispanic		−0.21** [−0.37, −0.05]		−0.11 [−0.28, 0.06]
AP gatekeeping × Asian/Pacific Islander		−0.11 [−0.25, 0.04]		−0.11 [−0.26, 0.03]
AP gatekeeping × American Indian		0.04 [−0.25, 0.33]		0.11 [−0.17, 0.40]
AP gatekeeping × two+ races		−0.07 [−0.22, 0.08]		−0.06 [−0.21, 0.09]
Prior achievement interactions				
AP gatekeeping × 9th-grade GPA at or above median (3.0)			—	
AP gatekeeping × 9th-grade GPA below median (3.0)			−0.37*** [−0.56, −0.17]	
AP gatekeeping × 9th-grade GPA (centered)				0.23 [−0.05, 0.51]
AP gatekeeping × 8th-grade math score (standardized)				0.03 [−0.03, 0.08]
AP gatekeeping × 8th-grade ELA score (standardized)				0.07 [−0.02, 0.16]
Covariates	X	X	X	X
School and cohort fixed effects	X	X	X	X
School observations	327	327	327	327
Student observations	144,669	144,669	144,669	144,669
Pseudo R^2	0.463	0.463	0.464	0.464

Source: Data are from the Wisconsin State Longitudinal Data Systems merged with survey data from the Office of Civil Rights Data Collection.
Note: Coefficients are from fixed-effects Poisson regressions. Complete model results are available in Table A.1 in the online supplement. The 95 percent confidence intervals (in brackets) are clustered at the school level. AP = Advanced Placement; ELA = English language arts.
+p < .1. *p < .05. **p < .01. ***p < .001 (two-tailed t test).

mechanisms for this pattern. It may be that the racial disproportionality in effects of gatekeeping is mainly explained by students' prior achievement; that is, gatekeeping only excludes racially minoritized students to the extent that those students have lower prior grades.

We partly test this proposition in Columns 3 and 4 in Table 3. First, in Column 3, we establish that gatekeeping mainly excludes students with relatively low prior grades by including an interaction term for below-median 9th-grade GPA students. Among at- or above-median GPA students, the estimated effect of gatekeeping is slightly negative and marginally statistically significant. However, the interaction is substantially negative and statistically significant such that gatekeeping reduces AP exam-taking by an estimated 41 percent $(1 - e^{-0.10-0.37} = 0.41)$ among students with lower than a B average. In Column 4, we estimate a model with interactions between gatekeeping and race/ethnicity and a set of interactions between gatekeeping and prior achievement. If the main mechanism for the racially disproportionate effects of gatekeeping is racial disparities in prior achievement, these interactions should account for much, if not all, of the racially disproportionate effects of the policy changes.[10] We find that the effect of AP gatekeeping on lower-achieving students does account for much of the nonsignificant racially disproportionate effects: The interaction term capturing the nonsignificant white-Black contrast in effects declines by about 80 percent, from -0.16 to -0.03, and the interaction term capturing the significant white-Hispanic contrast declines by about half, from -0.21 to -0.11, and becomes nonsignificant. This evidence suggests that lower average achievement among Hispanic and Black students compared to white students leads those students to be more affected by AP gatekeeping measures.

We next examine differences in the effects of gatekeeping among students with similar prior achievement as a further test of the importance of achievement in explaining the racially disproportionate effects of gatekeeping. Table 4 splits the sample into the top and bottom half of the GPA distribution and estimates our model of racial differences in gatekeeping effects for each subpopulation. We find that gatekeeping excludes Black and Hispanic students with below-median GPAs much more than it does white students with similar GPAs. The estimated effect among Asian/Pacific Islander students is also large but

not statistically significantly different from the effect among white students. In contrast, there are substantively smaller effects for all groups among students whose GPAs are above 3.0.[11] Because the magnitude of effects in percentage terms becomes more nonlinear as effects depart from zero, we visualize the differences by plotting the exponentiated coefficients and confidence intervals in Figure 3. Among white students and multiracial students, the estimated effect is to reduce exam-taking by near zero, compared to 56 percent among Black students, 45 percent among Hispanic students, 34 percent among Asian/Pacific Islander students, and 20 percent among American Indian/Alaska Native students.

In Appendices D and E in the online supplement, we conduct two additional analyses to probe these results. First, in Appendix D, we find that AP gatekeeping has similar effects on the number of AP exams students pass as the number they take. In Table E.1 in the online supplement, we explore the effects among below-median GPA students on exam-taking for specific AP subjects. We find similar overall effect sizes across subjects, albeit with somewhat smaller racial differences in English compared to math, STEM, and other non-STEM courses.

Effect of AP Enrollment Policies on the Transition to College

Finally, we examine whether gatekeeping policies ultimately affect racial inequality in postsecondary outcomes. If access to AP courses is important to students' propensity to attend selective colleges in particular, we expect to see racially disproportionate results in line with the effects on AP participation. Table 5 presents the results from separate logit models predicting two college outcomes: any four-year and highly selective four-year attendance, including interaction terms with race in the logit metric and group-specific average marginal effects (AMEs). We find suggestive evidence that overall college enrollment is slightly negatively affected by gatekeeping policies. However, there is clearer evidence that Black students are particularly less likely to attend four-year colleges in general, and highly selective four-year colleges in particular, under gatekeeping policies. The estimated effect of experiencing AP gatekeeping is to reduce Black students' odds of attending a four-year college by a factor of 0.48 $(e^{-0.13-0.60})$ and

Table 4. Estimated Effects of AP Gatekeeping Policy on Number of AP Exams Taken from fixed-Effects Poisson Models, Heterogeneous Effects by Race/Ethnicity and 9th-Grade GPA Quantile.

	Total number of AP exams taken	
	Below median GPA (3.0) (1)	At or above median GPA (3.0) (2)
Main effect		
AP gatekeeping	−0.01	−0.10*
	[−0.36, 0.34]	[−0.18, −0.01]
Race interactions		
AP gatekeeping × white	—	—
AP gatekeeping × Black	−0.80*	0.07
	[−1.57, −0.04]	[−0.18, 0.33]
AP gatekeeping × Hispanic	−0.60*	−0.07
	[−1.08, −0.12]	[−0.21, 0.07]
AP gatekeeping × Asian/ Pacific Islander	−0.42	−0.10
	[−1.14, 0.31]	[−0.23, 0.03]
AP gatekeeping × American Indian	−0.22	0.11
	[−1.09, 0.65]	[−0.21, 0.43]
AP gatekeeping × two+ races	−0.01	−0.06
	[−0.68, 0.65]	[−0.21, 0.10]
Covariates	X	X
School and cohort fixed effects	X	X
School observations	298[a]	326[a]
Student observations	70,979	72,358
Pseudo R^2	0.347	0.319

Source: Data are from the Wisconsin State Longitudinal Data Systems merged with survey data from the Office of Civil Rights Data Collection.
Note: Coefficients are from fixed-effects Poisson regressions. Complete model results are available in Table A.1 in the online supplement. The 95 percent confidence intervals (in brackets) are robust to heteroskedasticity and clustered at the school level. AP = Advanced Placement.
[a]The number of schools in these models is smaller than in the full sample because some schools have zero students taking AP courses within the subsamples split by median GPA. The fixed effects for those schools perfectly separate the outcome and must be dropped for estimation to be possible.
*$p < .05$.

a highly selective college by a factor of 0.31 ($e^{−0.23−0.95}$). The corresponding AMEs are reductions in the probability of enrollment by 7 and 2 percentage points, respectively. These effects are substantively significant and in all cases but the AME for elite college enrollment, statistically significant. We also find evidence of a large negative AME on Asian/Pacific Islander students' selective college enrollment of 5 percentage points. These findings suggest that when Black and Asian/Pacific Islander students gain access to AP courses through open enrollment policies, it facilitates their entry into baccalaureate colleges, highlighting the downstream consequences of policies that exacerbate racialized tracking.

Limitations

One limitation of this study is our survey-derived measure of schools' gatekeeping policies. There may be measurement error over time in schools' survey responses, perhaps driven by changes in the staff members filling out the OCRDC survey across years. If this is the case, the effects we estimate here are likely underestimates of the true effects of gatekeeping. However, we are
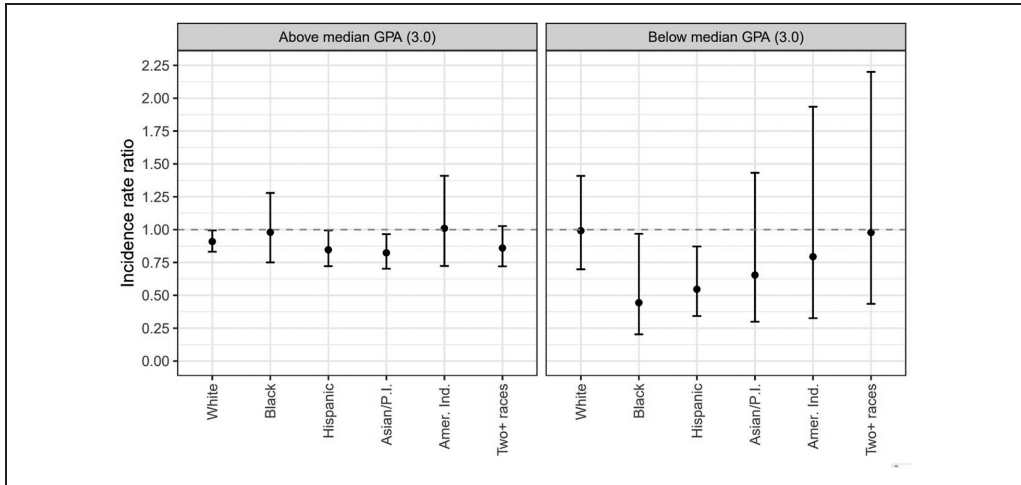
Figure 3. Estimated effect of gatekeeping policy on number of AP exams taken by race/ethnicity and ninth-grade GPA quantile.
Source: Data are drawn from the Wisconsin State Longitudinal Data Systems and the Office of Civil Rights Data Collection.
Note: Estimates and 95 percent confidence intervals in incidence rate ratio metric for each racial/ethnic group computed from Columns 1 and 2 in Table 4. AP = Advanced Placement.

encouraged by the results in Column 3 of Table 2, which show that our gatekeeping measure is strongly associated with the exclusion of lower-achieving students, just as we would expect if it was a strong and valid measure of policy.

Second, our analysis relies on the assumption that the timing of changes in a school's policies are random with respect to each new cohort's propensity to take AP courses, conditional on the covariates. We cannot be sure this assumption is met. Enrollment policy changes might coincide with other programs that encourage or discourage students, perhaps specifically students of color, from enrolling in AP courses. If this is the case, our estimates would instead identify the effect of the mix of policies that coincided. However, scholars have found that gatekeeping in AP courses is contentious among teachers and administrators (Rowland and Shircliffe 2016). We thus speculate that much of the variation we see in enrollment policies in schools over time reflects idiosyncratic turnover and decision-making processes among school staff.

Third, we do not directly observe what type of requirement—recommendations, grades, or another criterion—school officials enact. However, students' prior grades are probably central to any assignment process: Even teachers' decisions about recommendations likely hinge mainly

on prior performance. Furthermore, all types of requirements are similar in that they are unlikely to be applied strictly: School officials make exceptions to rules for certain students, for instance, when parents lobby for changes (Lewis and Diamond 2015). Relatedly, our measure of students' ninth-grade location in their school's tracking system—their classmates' test scores—is imperfect compared to a measure of specific prior coursework. Specific prior coursework levels are likely prerequisites for some AP courses, particularly in mathematics sequences, and students often face strong barriers to moving upward academically (Hanselman 2020; Kelly 2007; Kelly and Price 2011). However, our results are similar whether or not we include our measure of track location, suggesting our results would be robust to an even more precise measure.

Finally, we are limited to data from a single state. Wisconsin's unique qualities make it a compelling setting for answering these questions, but they may also limit the generalizability of our findings. Wisconsin is among the states with the most extreme racial inequality in both socioeconomic and academic outcomes (Smeeding and Thornton 2018; U.S. Department of Education 2022). And although the student population in the state has become more racially diverse, particularly as the Hispanic population has grown,

Table 5. Estimated Effects of AP Gatekeeping Policy on College Attendance from Fixed-Effects Logit Models, Heterogeneous Effects by Race/Ethnicity.

	Four-year college		Highly selective four-year college	
	Logit (1)	Group AME (2)	Logit (3)	Group AME (4)
Main effect				
AP gatekeeping	−0.13	—	−0.23	—
	[−0.44, 0.18]		[−0.54, 0.07]	
Race interactions				
AP gatekeeping × white	—	−0.02	—	−0.01
		[−0.07, 0.03]		[−0.03, 0.01]
AP gatekeeping × Black	−0.60***	−0.07**	−0.95**	−0.02
	[−0.90, −0.30]	[−0.11, −0.02]	[−1.58, −0.32]	[−0.05, 0.01]
AP gatekeeping × Hispanic	−0.01	−0.01	0.32	0.00
	[−0.23, 0.20]	[−0.05, 0.02]	[−0.24, 0.88]	[−0.02, 0.02]
AP gatekeeping × Asian/ Pacific Islander	−0.01	−0.02	−0.41^{+}	−0.05*
	[−0.31, 0.29]	[−0.08, 0.04]	[−0.90, 0.07]	[−0.10, 0.00]
AP gatekeeping × American Indian	−0.11	−0.02	−0.82	−0.02
	[−0.77, 0.55]	[−0.08, 0.04]	[−2.34, 0.71]	[−0.06, 0.01]
AP gatekeeping × two+ races	0.06	−0.01	−0.30	−0.02^{+}
	[−0.39, 0.52]	[−0.06 0.05]	[−0.79, 0.19]	[−0.06, 0.00]
Covariates	X		X	
School and cohort fixed effects	X		X	
School observations	326[a]		314[a]	
Student observations	144,664		143,417	
Pseudo R^2	0.357		0.397	

Source: Data are from the Wisconsin State Longitudinal Data Systems merged with survey data from the Office of Civil Rights Data Collection.
Note: Coefficients are from fixed-effects logistic regressions, presented in both logit metric and group-specific average marginal effects (AME). Complete model results are available in Table A.1 in the online supplement. The 95 percent confidence intervals (in brackets) are robust to heteroskedasticity and clustered at the school level. AP = Advanced Placement.[a]The number of schools in these models is smaller than in the full sample because a few schools have no students who attended a college of that type. The fixed effects for those schools perfectly separate the outcome and must be dropped for estimation to be possible.
$+p < .1$. *$p < .05$. **$p < .01$. ***$p < .001$ (two-tailed t test).

public school teachers in Wisconsin remain 96 percent white (Goff, Carl, and Yang 2018). This situation may heighten the racialized effects of gatekeeping if white teachers in Wisconsin are particularly likely to be biased against non-white students.

CONCLUSIONS

In this study, we found evidence that when teachers or other school officials enact selection criteria for advanced courses, they modestly reinforce racialized tracking and increase racial disparities in college enrollment. This finding is relevant to school and district policy: Removing gatekeeping measures may increase racial equity. Our results further suggest a synthesis of the scholarly literature on the potential effects of course gatekeeping on racialized tracking. First, our results support the quantitative literature emphasizing the importance of prior achievement disparities: Achievement criteria do disproportionately exclude any group with lower average achievement. However, we found this explanation alone wanting. AP gatekeeping disproportionately excludes Black and Hispanic

students relative to white students among those with similarly low prior achievement. This finding is consistent with the notion that teachers' expectations of students depend on their racial or ethnic background or that school officials are willing to override gatekeeping criteria for relatively low-achieving white students but unwilling to do so for relatively low-achieving Black and Hispanic students (Lewis and Diamond 2015). Alternatively, relatively low-achieving white students may be less likely than relatively low-achieving Black and Hispanic students to seek academic opportunities like AP courses. In practice, we cannot distinguish between these mechanisms with these data, but based on prior research, we suspect they both play a role in the results we report here.

Our results also partially accord with our first prediction in that they suggest course gatekeeping is but one, relatively minor mechanism through which racial dynamics and racism affect inequality in course enrollments. We found that gatekeeping largely affects racial inequality among a subset of lower-achieving students, suggesting a limited role for this dimension of organizational policies in the production of racialized tracking broadly. Supporting this conclusion, most schools in Wisconsin—and nationally (Farkas and Duffet 2009)—allow open enrollment in AP courses, yet there remain large racial disparities in participation. Other entrenched racialized and racist processes, including the formation of academic self-concept, segregated networks, and hostile academic environments, also reinforce racialized tracking. Furthermore, gatekeeping as we have conceptualized and measured it—formalized eligibility criteria—does not attend to teachers' or counselors' informal influence on students, for instance, by directly encouraging course enrollment based on their perception of students' academic performance.

Our findings suggest multiple avenues for future research. One fruitful avenue may be to continue examining the extent to which the effects of gatekeeping are due to student behavior as opposed to teacher or counselor bias, mechanisms we could not distinguish in our study (see Francis et al. 2019). The effects of gatekeeping may also vary depending on whether teachers use ''subjective'' or ''objective'' criteria (Grissom and Redding 2016). In the same vein, due to limited empirical variation in our context, we did not examine the role that teacher identity plays in these racialized processes despite the evidence for its importance in similar educational situations (e.g., Gershenson et al. 2016). Future work on this subject could integrate teacher characteristics to test this possibility. Finally, our findings on the postsecondary consequences of gatekeeping raise mechanistic questions. When students of color are induced to take AP courses, why do they appear to attend baccalaureate colleges at higher rates and at higher levels of selectivity? Does the effect reflect the admissions benefit of a more rigorous transcript, the cost-benefit shift associated with a potentially quicker college graduation timeline, a psychological impact of succeeding in advanced coursework, or something else? Future research, quantitative and qualitative, can investigate these pathways.

RESEARCH ETHICS

This study was approved by the Institutional Review Board at the authors' institution. The data used in this study are used in compliance with a data agreement with the Wisconsin Department of Public Instruction.

ORCID iDs

Noah Hirschl ⓘ https://orcid.org/0000-0003-0723-0733
Christian Michael Smith ⓘ https://orcid.org/0000-0002-8024-3832

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

NOTES

1. Our definition of course gatekeeping applies to earlier stages of educational stratification, including entry into gifted and talented programs. The concept also relates to academic criteria in later-stage transitions, such as the use of standardized test scores in selective college or graduate school admissions, although those decisions often involve many nonacademic criteria as well. In each of these cases, the type and stringency of performance-based admission criteria often have implications for racial equity. We thank an anonymous reviewer for making these connections.

2. This relationship may have additional complexity. In particular, course gatekeeping may influence the degree to which educational attitudes are racialized in a school. For example, gatekeeping may make attitudes more racialized by heightening the awareness among students of color that they must struggle for access to educational institutions, in turn building a pro-education attitude among students of color. Alternatively, gatekeeping may make attitudes less racialized by obstructing the efforts of students of color to take AP courses, in turn discouraging students of color and bringing down their attitudes closer to the attitudes of their similarly achieving white peers.

3. We begin tracking students in 10th grade because it is exceedingly rare for students to take AP exams in 9th grade: 1.4 percent of students did so in 2014, the first year of our study. We start in 10th grade rather than 11th because it has become reasonably common for 10th graders to take AP exams: 11 percent took at least one in 2016. We replicated our main models using only 11th- and 12th-grade AP exams as the outcome, and we found substantively similar, although slightly attenuated, results.

4. The National Student Clearinghouse data cover upward of 96 percent of all national postsecondary enrollment in higher-education institutions over the period we study (Dynarski, Hemelt, and Hyman 2015). We limit our measurement of enrollment to the year following high school to maintain equal measurement timing across the cohorts in the study.

5. The OCRCD survey is biennial, so we do not observe the policy measure in every school year. As a solution, we simply omitted the missing years when computing the average policy for each cohort of students (but not when computing each student's total number of AP exams or any other variable). As an alternative method, we linearly imputed the missing interim years for each school before averaging, and we found essentially the same results (available on request).

6. Given that the log-link functional form is correct, Poisson fixed-effects models yield consistent estimates of the conditional expectation function regardless of whether or not the outcome was generated under a Poisson process (see Wooldridge 1999). Yet analysts often model potential deviations from a Poisson process: overdispersion, often addressed with an overdispersion parameter, and excess zeros, often addressed with a zero-inflated or hurdle model. Our main results are substantively the same when estimated with an overdispersion parameter. And although our outcome data contain many zeros, we see no reason to assume that a separate process drives selection into taking zero AP courses, as would be assumed by a zero-inflated or hurdle model.

7. To construct this measure, we first calculate leave-one-out test score averages for each course a student is enrolled in: $\frac{1}{J-1}\left(\sum_{j \neq i} Test\ score_j\right)$, where J is the number of students enrolled in the course, each student j's $Test\ score$ is measured in z scores, and the focal student i's test score is left out of the mean calculation. We then take the average of these course-specific averages across all the courses the focal student is enrolled in to obtain the student's measure of track location.

8. The predominant reason students lack either GPA or test scores is because they were not enrolled in a Wisconsin public school in the grade in question. These students often entered the public school system in high school from a private school or from out of state. Thus, our results may only generalize to the majority of students who remained in the public school system from middle to high school.

9. We have adequate samples of students identifying with each racial subgroup at the schools that changed their policies, thus providing adequate statistical power to detect the effects of interest. Our sample of students at those schools includes 2,497 students identifying as Black, 4,199 identifying as Hispanic, 1,602 identifying as Asian or Pacific Islander, 530 identifying as American Indian, and 1,029 identifying with two or more races.

10. This proposition rests on the assumption of minimal measurement error in the linear combination of the academic achievement measures we use.

11. Any residual differences in achievement between racial groups *within* these two GPA quantiles could still drive differences in effects between racial groups. To test this possibility, in supplemental models (available on request), we replicated the model from Column 4 in Table 2 across our split sample by including three interaction terms between gatekeeping and prior GPA, eighth-grade math score, and eighth-grade reading score; the results were substantively the same. In Appendix C in the online supplement, we further test the possibility that the effects are driven by the clustering of some groups around the margin of AP participation.

REFERENCES

Alon, Sigal. 2009. ''The Evolution of Class Inequality in Higher Education: Competition, Exclusion, and Adaptation.'' *American Sociological Review* 74(5):731–55.

Avery, Christopher, Oded Gurantz, Michael Hurwitz, and Jonathan Smith. 2018. ''Shifting College Majors in Response to Advanced Placement Exam Scores.'' *Journal of Human Resources* 53(4):918–56.

Bastedo, Michael N., Joseph E. Howard, and Allyson Flaster. 2016. ''Holistic Admissions After Affirmative Action: Does 'Maximizing' the High School Curriculum Matter?'' *Educational Evaluation and Policy Analysis* 38(2):389–409.

Bennett, Pamela R., and Yu Xie. 2003. ''Revisiting Racial Differences in College Attendance: The Role of Historically Black Colleges and Universities.'' *American Sociological Review* 68(4):567–80.

Bowman, Nicholas A., and Michael N. Bastedo. 2018. ''What Role May Admissions Office Diversity and Practices Play in Equitable Decisions?'' *Research in Higher Education* 59(4):430–47.

Calarco, Jessica McCrory. 2020. ''Avoiding Us versus Them: How Schools' Dependence on Privileged 'Helicopter' Parents Influences Enforcement of Rules.'' *American Sociological Review* 85(2): 223–46.

Card, David, and Laura Giuliano. 2016. ''Universal Screening Increases the Representation of Low-Income and Minority Students in Gifted Education.'' *Proceedings of the National Academy of Sciences of the United States of America* 113(48): 13678–83.

Charles, Camille Z., Vincent J. Roscigno, and Kimberly C. Torres. 2007. ''Racial Inequality and College Attendance: The Mediating Role of Parental Investments.'' *Social Science Research* 36(1):329–52.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2003. ''Segregation and Resegregation in North Carolina's Public School Classrooms.'' *North Carolina Law Review* 81(4):1463–512.

College Board. 2019. ''2019 Program Summary Report.'' https://secure-media.collegeboard.org/digitalServices/pdf/research/2019/Program-Summary-Report-2019.pdf.

College Division of Barron's Profiles of American Colleges. 2008. *2009 Barron's Profiles of American Colleges*. Hauppauge, NY: Barron's Educational Series.

Conger, Dylan, Alec Kennedy, Mark C. Long, and Raymond McGhee, Jr. 2021. ''The Effect of Advanced Placement Science on Students' Skills, Confidence, and Stress.'' *Journal of Human Resources* 56(1): 93–124.

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(4):555–76.

Conwell, Jordan A. 2020. ''Diverging Disparities: Race, Parental Income, and Children's Math Scores, 1960-2009.'' *Sociology of Education* 110(9): 1689–99.

Cortes, Kalena E. 2010. ''Do Bans on Affirmative Action Hurt Minority Students? Evidence from the Texas Top 10% Plan.'' *Economics of Education Review* 29(6):1110–24.

Curtis, Katherine, and Sarah Lessem. 2014. *2010 Census Chartbook: Demographic Trends in Wisconsin*. Madison, WI: University of Wisconsin-Madison Applied Population Laboratory.

Diamond, John B. 2006. ''Still Separate and Unequal: Examining Race, Opportunity, and School Achievement in 'Integrated' Suburbs.'' *The Journal of Negro Education* 75(3):495–505.

Domina, Thurston, Andrew McEachin, Paul Hanselman, Priyanka Agarwal, NaYoung Hwang, and Ryan Lewis. 2019. ''Beyond Tracking and Detracking: The Dimensions of Organizational Differentiation in Schools.'' *Sociology of Education* 92(3):293–322.

Downey, Douglas B., James W. Ainsworth, and Zhenchao Qian. 2009. ''Rethinking the Attitude-Achievement Paradox among Blacks.'' *Sociology of Education* 82(1):1–19.

Dynarski, Susan M., Steven W. Hemelt, and Joshua M. Hyman. 2015. ''The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes.'' *Educational Evaluation and Policy Analysis* 37(1S):53S–79S.

Evans, Brent J. 2019. ''How College Students Use Advanced Placement Credit.'' *American Educational Research Journal* 56(3):925–54.

Farkas, Steve, and Ann Duffet. 2009. ''Growing Pains in the Advanced Placement Program: Do Tough Trade-Offs Lie Ahead?'' The Thomas B. Fordham Institute. Retrieved December 2022. https://fordhaminstitute.org/national/research/growing-pains-advanced-placement-program-do-tough-trade-offs-lie-ahead.

Ferguson, Ronald F. 2003. ''Teachers' Perceptions and Expectations and the Black-white Test Score Gap.'' *Urban Education* 38(4):460–507.

Fox, Lindsay. 2016. ''Seeing Potential: The Effects of Student–Teacher Demographic Congruence on Teacher Expectations and Recommendations.'' *AERA Open* 2(1). doi:10.1177/2332858415623758.

Francis, Dania V., and William A. Darity. 2020. ''Isolation: An Alternative to the 'Acting white' Hypothesis in Explaining Black Under-Enrollment in Advanced Courses.'' *Journal of Economics, Race, and Policy* 3(2):117–22.

Francis, Dania V., Angela C. M. De Oliveira, and Carey Dimmitt. 2019. ''Do School Counselors Exhibit Bias in Recommending Students for Advanced Coursework?'' *B.E. Journal of Economic Analysis and Policy* 19(4):1–47.

Gamoran, Adam. 1992. ''The Variable Effects of High School Tracking.'' *American Sociological Review* 57(6):812–28.

Gamoran, Adam, and Robert D. Mare. 1989. ''Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?'' *American Journal of Sociology* 94(5):1146–83.

Gershenson, Seth, Stephen B. Holt, and Nicholas W. Papageorge. 2016. ''Who Believes in Me? The Effect of Student–Teacher Demographic Match on Teacher Expectations.'' *Economics of Education Review* 52:209–24.

Goff, Peter, Bradley Carl, and Minseok Yang. 2018. ''Supply and Demand for Public School Teachers in Wisconsin.'' WCER Working Paper No. 2018-2. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin–Madison.

Grissom, Jason A., and Christopher Redding. 2016. ''Discretion and Disproportionality: Explaining the Underrepresentation of High-Achieving Students of Color in Gifted Programs.'' *AERA Open* 2(1). doi: 10.1177/2332858415622175.

Gurantz, Oded. 2019. ''How College Credit in High School Impacts Postsecondary Course-Taking: The Role of AP Exams.'' EdWorkingPaper No. 19-110. Providence, RI: Annenberg Institute at Brown University.

Hanselman, Paul. 2020. ''Tournament Mobility in Mathematics Course-Taking Pathways.'' *Socius* 6. doi:10.1177/2378023120927604.

Hout, Michael, and Thomas A. DiPrete. 2006. ''What We Have Learned: RC28's Contributions to Knowledge about Social Stratification.'' *Research in Social Stratification and Mobility* 24(1):1–20.

Karlson, Kristian Bernt. 2015. ''Expectations on Track? High School Tracking and Adolescent Educational Expectations.'' *Social Forces* 94(1):115–41.

Kelly, Sean. 2007. ''The Contours of Tracking in North Carolina.'' *The High School Journal* 90(4):15–31.

Kelly, Sean. 2009. ''The Black-white Gap in Mathematics Course Taking Effects of Mathematics Course.'' *Sociology of Education* 82(1):47–69.

Kelly, Sean, and William Carbonaro. 2012. ''Curriculum Tracking and Teacher Expectations: Evidence from Discrepant Course Taking Models.'' *Social Psychology of Education* 15(3):271–94.

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(3): 560–85.

Kolluri, Suneal. 2018. ''Advanced Placement: The Dual Challenge of Equal Access and Effectiveness.'' *Review of Educational Research* 88(5):671–711.

Lewis, Amanda E., and John B. Diamond. 2015. *Despite the Best Intentions: How Racial Inequality Thrives in Good Schools*. New York, NY: Oxford University Press.

Lewis-McCoy, R. L'Heureux. 2014. *Inequality in the Promised Land: Race, Resources, and Suburban Schooling*. Stanford, CA: Stanford University Press.

Lucas, Samuel. 1999. *Tracking Inequality: Stratification and Mobility in American High Schools*. New York, NY: Teachers College Press.

Lucas, Samuel R., and Mark Berends. 2002. ''Sociodemographic Diversity, Correlated Achievement, and De Facto Tracking.'' *Sociology of Education* 75(4): 328–48.

Lucas, Samuel R., and Mark Berends. 2007. ''Race and Track Location in U.S. Public Schools.'' *Research in Social Stratification and Mobility* 25(3):169–87.

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* 46:355–78.

Malkus, Nat. 2016. ''AP at Scale: Public School Students in Advanced Placement, 1990–2013.'' American Enterprise Institute Report. https://www.aei.org/research-products/report/ap-at-scale-public-school-students-in-advanced-placement-1990-2013/.

Melguizo, Tatiana. 2010. ''Are Students of Color More Likely to Graduate from College If They Attend More Selective Institutions? Evidence from a Cohort of Recipients and Nonrecipients of the Gates Millennium Scholarship Program.'' *Educational Evaluation and Policy Analysis* 32(2):230–48.

Michelmore, Katherine, and Susan M. Dynarski. 2017. ''The Gap within the Gap: Using Longitudinal Data to Understand Income Differences in Educational Outcomes.'' *AERA Open* 3(1). doi:10.1177/233285841769.

Mickelson, Roslyn Arlin. 1990. ''The Attitude-Achievement Paradox among Black Adolescents.'' *Sociology of Education* 63(1):44–61.

Mickelson, Roslyn Arlin. 2001. ''Subverting Swann: First- and Second-Generation Segregation in the Charlotte-Mecklenburg Schools.'' *American Educational Research Journal* 38(2):215–52.

Milewski, Glenn B., and Jacqueline M. Gillie. 2002. ''What Are the Characteristics of AP® Teachers? An Examination of Survey Research.'' The College Board Research Report No. 2002-10. College Entrance Examination Board, New York, NY.

Moody, James. 2001. ''Race, School Integration, and Friendship Segregation in America.'' *American Journal of Sociology* 107(3):679–716.

Oakes, Jeannie. 1994. ''More Than Misapplied Technology: A Normative and Political Response to Hallinan on Tracking.'' *Sociology of Education* 67(2):84–91.

O'Connor, Carla, Jennifer Mueller, R. L'Heureux Lewis-McCoy, Deborah Rivas-Drake, and Seneca Rosenberg. 2011. '''Being' Black and Strategizing for Excellence in a Racially Stratified Academic Hierarchy.'' *American Educational Research Journal* 48(6):1232–57.

Price, Heather E. 2021. "The College Preparatory Pipeline: Disparate Stages in Academic Opportunities." *American Educational Research Journal* 58(4):785–814.

Ray, Victor. 2019. "A Theory of Racialized Organizations." *American Sociological Review* 84(1):26–53.

Rodriguez, Awilda, and Keon M. McGuire. 2019. "More Classes, More Access? Understanding the Effects of Course Offerings on Black-white Gaps in Advanced Placement Course-Taking." *The Review of Higher Education* 42(2):641–79.

Rosenbaum, James E. 1976. *Making Inequality: The Hidden Curriculum of High School Tracking.* New York, NY: John Wiley & Sons Inc.

Rowland, M. Lance, and Barbara J. Shircliffe. 2016. "Confronting the 'Acid Test': Educators' Perspectives on Expanding Access to Advanced Placement at a Diverse Florida High School." *Peabody Journal of Education* 91(3):404–20.

Schaeffer, Katherine. 2021. "America's Public School Teachers Are Far Less Racially and Ethnically Diverse Than Their Students." Pew Research Center. https://www.pewresearch.org/fact-tank/2021/12/10/americas-public-school-teachers-are-far-less-racially-and-ethnically-diverse-than-their-students/.

Schneider, Jack. 2009. "Privilege, Equity, and the Advanced Placement Program: Tug of War." *Journal of Curriculum Studies* 41(6):813–31.

Small, Mario L., and Christopher Winship. 2007. "Black Students' Graduation from Elite Colleges: Institutional Characteristics and between-Institution Differences." *Social Science Research* 36(3):1257–75.

Smeeding, Timothy M., and Katherine Thornton. 2018. *Poverty, Incomes, Race and Ethnicity in Wisconsin and Milwaukee: A Supplement to the 2016 Wisconsin Poverty Report.* Madison, WI: Institute for Research on Poverty.

Smith, Jonathan, Michael Hurwitz, and Christopher Avery. 2017. "Giving College Credit Where It Is Due: Advanced Placement Exam Scores and College Outcomes." *Journal of Labor Economics* 35(1):67–147.

Sørensen, Aage Bøttger. 1970. "Organizational Differentiation of Students and Educational Opportunity." *Sociology of Education* 43(4):355–76.

Spade, Joan Z., Lynn Columba, and Beth E. Vanfossen. 1997. "Tracking in Mathematics and Science: Courses and Course-Selection Procedures." *Sociology of Education* 70(2):108–27.

Spiegelman, Maura. 2020. "Race and Ethnicity of Public School Teachers and Their Students." NCES 2020-103. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020103.

Taliaferro, Jocelyn De Vance, and Jessica T. DeCuir-Gunby. 2008. "African American Educators' Perspectives on the Advanced Placement Opportunity Gap." *Urban Review* 40(2):164–85.

Tyson, Karolyn. 2011. *Integration Interrupted: Tracking, Black Students, and Acting white after Brown.* New York, NY: Oxford University Press.

U.S. Department of Education. 2022. Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2022 Reading Assessment. Washington, DC: U.S. Department of Education. https://www.nationsreportcard.gov/.

Wooldridge, Jeffrey M. 1999. "Distribution-Free Estimation of Some Nonlinear Panel Data Models." *Journal of Econometrics* 90(1):77–97.

Yonezawa, Susan, Amy Stuart Wells, and Irene Serna. 2002. "Choosing Tracks:'Freedom of Choice' in Detracking Schools." *American Educational Research Journal* 39(1):37–67.

Author Biographies

Noah Hirschl is a PhD candidate in sociology at the University of Wisconsin-Madison. His research focuses on postsecondary education and social stratification. His current research projects examine inequality among graduate and professional degrees, racial and ethnic differences in the process of educational attainment, and educational assortative mating.

Christian Michael Smith is a postdoctoral fellow at the University of California, Merced. His research seeks to identify solutions to pernicious stratification, especially socioeconomic inequalities in U.S. higher education.