Which Match Matters? HBCUs, Same-Race Teacher Effects, and Black Student Achievement*

Lavar Edmonds
Stanford University

This Draft: January 2026

Abstract

Researchers study racial concordance effects in many settings, but why matching matters remains unclear. I test channels by studying the impacts of teachers trained at Historically Black Colleges and Universities (HBCUs) on Black student outcomes in North Carolina elementary schools. Assigning an HBCU-trained teacher raises Black students' end-of-grade math exam scores and reduces suspensions. Both Black and White HBCU-trained teachers outperform same-race, non-HBCU peers, indicating race-independent gains. Results are consistent with treatment effects – different teacher education practices – rather than selection into HBCUs. Together, this work implies a role for training and learning in affecting outcomes beyond racial (mis)match. JEL Codes: I21, I23, I24, J15

*Contact: ledmonds@stanford.edu. A previous version of this manuscript was circulated under the title, "Role Models Revisited: HBCUs, Same-Race Teacher Effects, and Black Student Achievement." Thanks to Kara Bonneau and the North Carolina Education Research Data Center for providing data and administrative assistance. Special thanks to Jennifer Jennings for both her research guidance and her generous purchase of the NCERDC data. Likewise, I am grateful to Leah Boustan, Jesse Bruhn, Brian Cadena, Janet Carlson, Tom Dee, Ellora Derenoncourt, Matthew Desmond, Michael Gottfried, Peter Hinrichs, Christopher Neilson, Isaac Sorkin, Karolyn Tyson, Jon Valant, Anne Kat Alexander, Joe Fish, Robert Kelchen, Jennifer Kowalski, Lillian Leung, Jacob Light, Renee Louis, Brendan Moore, João Pugliese, and China Mae Stepter for insightful comments and suggestions on previous proposals and drafts of the paper. I also thank seminar and conference participants from the Princeton Education Research Section, the Eviction Lab, the Economics of Racism, the Urban Institute, the NBER Race and Stratification working group, and the W.E. Upjohn Institute for Employment Research for their helpful feedback. All errors – and there will be some – are all my own.

1 Introduction

Extensive research across the social sciences explores same-race match effects, including studies of mentors and impacts on occupational choice and job satisfaction, judges and legal proceedings, police-civilian interactions and criminal justice reform, physician-patient matches and health outcomes, and even basketball referees and game penalties.[1] Even papers without an explicit focus on racial match have relied on assumptions about same-race pairings and behavior between individuals.[2] Education research has similarly proposed that students benefit from exposure to demographically-congruent teachers, especially racial and ethnic minoritized students.[3] However, this work on racial representation focuses on whether race matters for outcomes without demonstrating how or why race matters, leaving the match effect mechanisms largely unspecified. Determining what underlies potential returns to same-race pairings is policy relevant given the substantive difference between inherent, immutable characteristics that same-race matches share and a bundle of skills that can be learned.

In this paper, I revisit same-race match effects and their mechanisms through an analysis of teachers trained at Historically Black Colleges and Universities (HBCUs). Specifically, I address two questions. First, what is the impact of having an HBCU-trained teacher on Black student outcomes? Second, how does HBCU attendance interact with same-race teacher effects for Black students? HBCUs are an appropriate context to study (racial) match effects given their historical and contemporary significance in the education of Black teachers, with approximately 50% of Black teachers having attended an HBCU (National Association for Equal Opportunity in Higher Education, 2008). They also train a considerable number of White teachers – in this sample, nearly one-quarter of HBCU-trained teachers are White – providing an observable environment where trainees of different races are exposed to race-conscious pedagogical practices, and potentially learned skills can be measured after their matriculation into K-12 classrooms.

I study HBCU attendance and teacher quality by estimating teacher effects on end-

[1] Abrams, Bertrand, and Mullainathan, 2012; Alsan, Garrick, and Graziani, 2019; Ba et al., 2021; Greenwood et al., 2020; Grissom and Keiser, 2011; Headley and Wright, 2020; Hoekstra and Sloan, 2022; Kofoed and McGovney, 2019; Price and Wolfers, 2010; Steffensmeier and Britt, 2001.

[2] Arnold, Dobbie, and Yang, 2018; Rose, 2021.

[3] Bartanen and Grissom, 2021; Dee, 2004, 2005; Delhommer, 2022; Egalite, Kisida, and Winters, 2015; Gershenson et al., 2022; Holt and Gershenson, 2019; Klopfenstein, 2005; Penney, 2017; Villegas and Lucas, 2004.

of-grade math and English Language Arts (ELA) exams using administrative panel data from elementary schools in North Carolina, producing four important results. First, I find HBCU-trained teachers have a positive effect on Black student achievement. Different specifications indicate that, on average, Black students score 0.026-0.041 standard deviations (SD) higher on standardized math exams in the year they are assigned to an HBCU-trained teacher. Effects are substantively large: Black student assignment to HBCU-trained teachers reduces their within-school Black-White test score gap by roughly 20%. Estimated magnitudes are at least comparable to those reported in previous papers on same-race teacher matches. Results are robust to alternative estimators, sample restrictions, and rich student and teacher controls. Notably, I do not observe significant negative effects for non-Black students. I also present results from corresponding analyses for ELA test scores. As with prior studies that show null teacher effects on reading scores, I find positive but small effects (0.01 SD) that are more sensitive to model specification.

Second, I only find a same-race match effect on standardized exams for Black students when the teacher attended an HBCU. In fact, Black students score higher on math exams when paired with Black HBCU-trained teachers over Black teachers from non-HBCUs. This within-race result extends to Black students matched with White HBCU-trained teachers (compared against White teachers from other colleges), implying the existence of race-independent teacher characteristics underlying previously observed match effects that improve student outcomes.

Third, I show HBCU-trained teachers have meaningful effects on lowering Black student suspensions, especially for Black boys. Additionally, whereas there are no significant differences in suspension probabilities between HBCU-trained and non-HBCU-trained Black teachers, White teachers from HBCUs are appreciably less likely to suspend Black students than their White counterparts from non-HBCUs.

Finally, I present evidence that these results reflect treatment effects – pedagogical training and skills developed at HBCUs – rather than selection of certain types of students into HBCUs. First, using national survey data on first-time college freshmen, I demonstrate that a college's distance from a student's home is the predominant factor influencing attendance at HBCUs. Proximity is especially salient for White HBCU attendees and is predictive above and beyond variables that capture students' pre-college orientation towards civic engagement and social justice, allaying concerns that students potentially select into HBCUs on unobservable characteristics that correlate with teacher effectiveness with Black students. Second, using college major and

teacher certification data available for over one-third of the teachers in the North Carolina sample, I show that HBCU-trained teachers with majors in education-related fields are more effective with Black students than their peer education graduates from other institutions. I then corroborate these results using data from the aforementioned national survey that links responses of students from the aforementioned freshmen survey to the follow-up survey administered to these same students as college seniors. This unique data set enables me to explore how student attitudes change over the course of attending different types of colleges. Comparing across students' senior year responses, students at HBCUs exhibit greater orientation towards civic and racial justice than their peers do, even conditional on entering college with the same inclination levels. Moreover, when comparing students from different colleges but with the same majors, only education majors at HBCUs demonstrate significant increases in civic and racial justice orientation over their counterparts.

In summary, this paper makes two key contributions to prior literature. Principally, this research reframes same-race effects and suggests that positive outcomes for same-race-matched individuals reflect a bundle of (learnable) skills, rather than inherent attributes determined by race. It also presents higher education training as a scalable policy path for how such skills may be acquired and developed by different-race individuals. If these relevant, human capital-boosting skills can be replicated at other training institutions (e.g., medical school, law school, etc.), my findings pose a promising area for future practitioner training and development policies. Specific to education research and policy, this paper also advances literatures on teacher preparation in higher education, teacher value-added, and teacher recruitment and retention.[4]

This paper is structured as follows: in Section 2, I briefly discuss the history and relevance of Historically Black Colleges and Universities, especially as they pertain to Black teachers and Black pedagogical training. Section 3 outlines the data used, and Section 4 details the econometric models for the analysis. Section 5 presents results on the effects of HBCU-trained teachers on Black student achievement. I show results are robust to alternative specifications and estimators, and I build on these findings to identify variation in Black student achievement conditional on both the teacher's HBCU attendance and race. I highlight and empirically test possible mechanisms for these results in Section 6 before concluding and discussing policy implications in

[4]Bacher-Hicks and Koedel, 2023; Bartanen and Grissom, 2021; Goldhaber et al., 2013; Ingersoll et al., 2017; Jackson et al., 2014; Koedel et al., 2015; Kumar and Lauermann, 2018.

Section 7.

2 Historically Black Colleges and Universities in the United States

Historically Black Colleges and Universities have been staples of Black communities and avenues for Black student access to post-secondary education for nearly two centuries (Harper and Mawhinney, 2017). Largely clustered in the U.S. South, but spanning states from Pennsylvania to Florida, there are roughly 100 Historically Black Colleges and Universities operating today in 19 states, D.C., and the U.S. Virgin Islands. These colleges have traditionally served majority Black student populations, though HBCUs specifically refer to colleges founded before 1964 in response to the *de jure* and *de facto* exclusion of Black people from higher education. As such, each was founded with the broad mission of educating Black Americans (Roebuck and Murty, 1993, pg. 3).

Although they are important institutions for Black college graduates across all fields, HBCUs historically have been especially crucial to the supply of Black teachers in the United States. Many HBCUs were founded, or at least existed at some point, as "normal schools," institutions created to prepare teachers for the classroom (Akbar and Sims, 2008). Today, HBCUs carry the legacy of an era where Black teachers, principals, and colleges worked explicitly in concert to develop curriculum motivated by and geared directly towards Black students. As detailed in Acosta, Foster, and Houchen (2018) and Givens (2021), the history of Black pedagogical praxis and expertise comprises decades of intentional teaching practices cultivated and circulated by a rich network of national, state, and local associations of Black teachers and principals via conferences and publications. Moreover, K-12 schools' connections to HBCUs facilitated a pipeline for recruiting promising Black high school students into teaching, who were trained at these colleges and in turn able to use those relationships to obtain employment in K-12 schools following graduation.

Remarkably, HBCUs' contributions to the Black teacher supply continued even after the Civil Rights Act of 1964's expansion of collegiate opportunities for Black students at non-HBCUs. Based on 2019-2020 degree conferral data from the Integrated Postsecondary Education Data System (IPEDS), while HBCUs produce only about 8% of Black college graduates, they graduated just over 20% of the Black college graduates with education degrees, suggesting HBCUs remain a pivotal desti-

nation for Black students interested in teaching [5] Further, previous estimates report HBCUs have produced roughly 50% of the current supply of Black teachers in the U.S. (National Association for Equal Opportunity in Higher Education, 2008).

HBCUs, long central to training Black teachers, remain largely unexplored in economics. Existing work centers on finance and funding (Sav, 2010), accountability and employment (Price, Spriggs, & Swinton, 2011), wages (Mykerezi & Mills, 2008), and educational outcomes (Edwards et al., 2025; Wilson, 2007). To my knowledge, no published study examines HBCUs' teacher-preparation programs or the effectiveness of their graduates. This paper fills that gap by providing the first evidence on HBCU-trained teachers' effects on Black students' outcomes. Prior research suggests preservice preparation can shape teachers' practice, especially with racially/ethnically minoritized students (e.g., Kumar & Lauermann, 2018), and debates whether program quality and type affect student outcomes (e.g., Goldhaber, Liddle, & Theobald, 2013; Koedel et al., 2015). I extend that literature by focusing explicitly on teacher preparation at HBCUs and its direct impact on student academic performance.

2.1 Teacher training as an alternative explanation for same-race teacher effects

Teacher training reveals a new lever for understanding same-race teacher effects beyond those previously discussed in economics research, such as role model effects, stereotype threat, and teacher expectations. In the first case, Black teacher success with Black students may stem from exposure to educated, successful Black adults in the classroom, updating their beliefs on returns to their own academic achievement.[6] In the second, same-race teachers may alleviate the student's anxiety of having their performance be an indicator of and benchmark for their racial group's (negatively) perceived intellect. Finally, different-race teachers with lower academic expectations of their students and implicit biases may adversely influence student effort and educational attainment.

While these theories may suffice in some settings, ultimately they fall short in fully explaining same-race teacher effects. Extant same-race teacher research, as with the broader same-race match literature, poses little evidence on underlying mechanisms, in part because research often relies on race alone as the defining feature. Being Black

[5] Based on the author's calculations.

[6] For further discussion and critiques of the role model effects hypothesis, see Harper and Davis (2012), Irvine (1989, 1990), Maylor (2009), and Villegas and Lucas (2004).

does not automatically enable teachers to positively affect Black student outcomes. Even if a teacher possesses cultural knowledge for basing instruction, there is nothing innate about race that guarantees the capacity and language to do so.[7] Alternatively, teachers require some foundation, like formal teacher education, for guidance in instruction. As such, a more suitable framework for understanding same-race teacher effects would account for diversity in teacher training and perspective.

3 Data

I draw on administrative data for 3rd, 4th, and 5th grade students and teachers in public schools hosted by the North Carolina Education Research Data Center (NCERDC). Students and teachers in the data are assigned unique IDs that allow researchers to track them across multiple years and multiple data sets. These data contain a host of student-level demographic, socioeconomic, and academic characteristics, such as race, gender, economic disadvantage status, and academic (test scores) and social-behavioral (suspensions, daily attendance) outcomes, as well as demographic counts for the schools they attend. Teacher data include demographic characteristics (race, gender, and age), pre-service history (colleges attended, level of education, certification test scores), and years of experience teaching in North Carolina.

As in prior studies using North Carolina education data (e.g., Bates et al., 2023; Hill and Jones, 2021; Jackson, 2018), I use a course membership file to link students to teachers who taught specific courses during a year. For more credible estimates, I impose several sample restrictions, including dropping non-traditional schools (e.g., charters, magnets, alternative schools), restricting to students with at least two years of test scores, removing classrooms with fewer than 10 or greater than 50 students, and removing classrooms where greater than 50% of students are designated recipients of some special accommodations (i.e., gifted, disabilities, or English language learners). In turn, I compile a sample of 247,286 unique Black students (and 1,091,721 unique students total) taught by 27,501 unique teachers over the 11-year period, covering 539,423 Black student-year observations in math and ELA.[8]

[7] Other scholars have made a similar argument. Cherry-McDaniel (2019) cautions against assuming "teachers of color are inherently capable of teaching in culturally responsive and sustaining ways" and posits that that assumption "has led to an overall uncritical approach to diversity training in teacher preparation programs" (pgs. 249-250).

[8] See the Data Appendix for additional details on sample construction and the matching procedure.

3.1 Summary statistics for student and teacher samples

Table 1. Summary statistics for teachers, 2008–2018

	(1) HBCU-trained	(2) Non-HBCU-trained	(1)-(2) Difference
Panel A: All teachers			
Black	73.92	6.33	67.59***
White	23.94	90.17	-66.23***
Hispanic	0.87	1.02	-0.15
Asian	0.25	0.59	-0.34***
Female	92.74	93.16	-0.43
Mean age	43.26	38.99	4.27***
Mean years of experience	14.14	11.58	2.56***
Graduate degree	44.04	39.03	5.01***
Graduate degree from an HBCU	14.86	1.81	13.05***
Mean NTE/Praxis z score	-0.52	0.23	-0.75***
Attended college out of state	10.69	29.81	-19.12***
National Board Certified	7.30	12.82	-5.52***
Mean classes taught in year	1.53	1.42	0.11***
Mean number of Black students taught in a year	14.30	7.48	6.82***
Mean share of Black students in school	46.98	24.77	22.21***
Panel B: Black teachers			
Female	92.63	92.38	0.25
Mean age	43.36	39.45	3.91***
Mean years of experience	15.17	10.55	4.62***
Graduate degree	49.58	50.98	-1.40
Graduate degree from an HBCU	16.75	6.18	10.58***
Mean NTE/Praxis z score	-0.73	-0.29	-0.44***
Attended college out of state	11.61	31.57	-19.96***
National Board Certified	5.75	6.11	-0.36
Mean classes taught in year	1.51	1.43	0.08***
Mean number of Black students taught in a year	15.81	14.87	0.94**
Mean share of Black students in school	51.99	48.01	3.98***
Panel C: White teachers			
Female	92.76	93.26	-0.50
Mean age	43.44	39.03	4.41***
Mean years of experience	11.48	11.73	-0.25
Graduate degree	28.62	38.23	-9.61***
Graduate degree from an HBCU	9.83	1.52	8.31***
Mean NTE/Praxis z score	0.10	0.28	-0.18***
Attended college out of state	8.28	29.70	-21.42***
National Board Certified	12.24	13.57	-1.33
Mean classes taught in year	1.58	1.41	0.16***
Mean number of Black students taught in a year	9.93	6.92	3.00***
Mean share of Black students in school	31.97	22.98	8.98***
Observations (Panel A)	2,423	25,078	
Observations (Panel B)	1,791	1,587	
Observations (Panel C)	580	22,613	

Notes: Welch unequal-variance two-sample t-tests are used for all differences. Shares (Black, White, Hispanic, Asian, Female, Graduate degree, Graduate degree from an HBCU, Attended college out of state, National Board Certified) are displayed in percentage points. Time-invariant characteristics are taken at the teacher level. Time-varying measures are averaged across years per teacher. Stars denote significance: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

HBCU-trained and non-HBCU-trained teachers differ considerably along virtually all observable characteristics.[9] Regressions will include indicators for whether a teacher obtained a graduate degree from an HBCU. Following Table 1, demographically, the two groups largely split along race, with mostly Black HBCU-trained teachers and mostly White non-HBCU-trained teachers (though almost one-fourth of HBCU-trained teachers are White). HBCU-trained teachers are far more likely to teach in schools with a greater proportion of Black students and have more experience teaching Black students. Non-HBCU-trained teachers in this sample more often than their HBCU-trained peers attended a college outside of North Carolina, an important difference if students at HBCUs are systematically more likely to receive training closely aligned with North Carolina's curricular standards. Lastly, a large gap exists in teacher certification test scores for HBCU-trained and non-HBCU-trained teachers. Differences between HBCU and non-HBCU attendees are largely preserved when looking within teacher racial groups.

I also display counts of teachers from the 10 largest teaching-producing colleges in the sample (Appendix Table 1). The teacher pipeline is overwhelmingly in-state and somewhat concentrated in a relatively small group of colleges. Most non-HBCU teachers come from medium-to-large public universities, especially the UNC system, representing 1,129 distinct colleges. That said, Appalachian State and East Carolina alone supply over 18% of all non-HBCU-trained teachers. On the HBCU side, I observe graduates from every North Carolina HBCU (11 total) plus 53 out-of-state HBCUs. White teachers trained at HBCUs are notably concentrated: Although they arrive from 24 unique HBCUs, four North Carolina campuses - Fayetteville State, Elizabeth City State, North Carolina Central, and Winston-Salem State — account for 77% of that group. Black HBCU-trained teachers are also concentrated, albeit relatively more dispersed (they come from 60 HBCUs and have greater representation from all HBCUs in the state).

[9]Throughout this paper, I use "HBCU-trained" to refer to teachers who attended an HBCU for their undergraduate education and "non-HBCU-trained" to refer to teachers who attended a non-HBCU for their undergraduate education.

Table 2. Summary statistics for students, 2008–2018

	(1) Group A	(2) Group B	(3) (1)-(2)
Panel A: Pre-treatment (t-1) among ever-treated vs. first observed year among never-treated			
Female	0.496	0.496	0.000
Economically disadvantaged	0.804	0.771	0.033***
Disability	0.115	0.141	-0.026***
LEP	0.006	0.008	-0.002***
Suspended	0.145	0.128	0.017***
Days absent	5.463	5.362	0.101***
Math test score	-0.532	-0.471	-0.060***
ELA test score	-0.453	-0.413	-0.041***
Panel B: Ever-treated vs never-treated			
Female	0.497	0.496	0.001
Ever economically disadvantaged	0.859	0.820	0.039***
Ever disability	0.152	0.170	-0.018***
Ever LEP	0.007	0.009	-0.002***
Ever suspended	0.278	0.254	0.024***
Avg. days absent	5.710	5.496	0.214***
Avg. math test score	-0.534	-0.470	-0.063***
Avg. ELA test score	-0.471	-0.422	-0.049***
Observations (Panel A)	28,895	171,343	
Observations (Panel B)	75,943	171,343	

Notes: Group A refers to students ever assigned to an HBCU-trained teacher. Group B refers to students never assigned to an HBCU-trained teacher. Welch unequal-variance two-sample t-tests are used for all differences. Panel A compares characteristics measured in t-1 among students who first meet an HBCU-trained teacher at year t to the first observed year among students who are never exposed; students whose first observed year is an HBCU year and students always exposed to HBCU teachers are excluded. Panel B reports student-level averages across all observed years (both groups). Stars denote significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Black students ever assigned to at least one HBCU-trained teacher and those who never have one statistically differ on most observables (Table 2), though there is also some evidence of negative selection into HBCU-trained teacher classrooms. These students have considerably lower average test scores, are more likely to have received a suspension judgment, and more often from economically disadvantaged backgrounds.

3.2 Comparing HBCUs in North Carolina and nationally

Based on tabulations of degree completion data from the Integrated Postsecondary Education Data System (IPEDS) (Appendix Table 2), North Carolina HBCUs look broadly similar to HBCUs nationwide, with the main difference being a smaller share of non-Black/non-White students. Within education departments, NC HBCUs are less Black and more White than the national HBCU average, though these gaps are statistically insignificant. Nationally, education is a more popular major at HBCUs for both Black and White students than at non-HBCUs; in North Carolina, the advantage is more muted. Notably, education programs at NC HBCUs are more racially integrated than the institutions as a whole: White students account for about 11% of all NC-HBCU graduates, but make up roughly one-third of education graduates. This composition helps explain the sizable presence of White HBCU-trained teachers in the NCERDC sample.

4 Empirical Strategy

4.1 Modeling student achievement

I estimate teacher effects using two linear and additively separable regression specifications akin to the one-step value-added model commonly used in education research (Bacher-Hicks and Koedel, 2023). First, I employ a two-way fixed effects estimator (TWFE) that relies on within-student variation in teacher assignment over time. Potential endogeneity arises from the non-random assignment of students to teachers (Paufler and Amrein-Beardsley, 2014), which could complicate teacher quality identification if teachers are assigned systematically different students (e.g., higher scoring) from ostensibly less-effective teachers. Within-student estimation strengthens empirical claims by eliminating time-invariant student characteristics (e.g., student ability) that may otherwise bias teacher effectiveness estimates. The corresponding regression equation is presented in Equation (1).

$$A_{it} = \beta(\text{HBCU}_{it}) + X'_{ijst}\gamma + \delta_i + \eta_t + \epsilon_{it}. \tag{1}$$

The dependent variable is the score on an end-of-grade exam (with separate equations estimated for math and ELA), standardized within grade and year to mean 0 and standard deviation 1, for a student i in year t. I regress this test score on a vector X of time-varying student, teacher (j), and school (s) characteristics, as well as student (δ_i)

and year (η_t) fixed effects. The coefficient of interest is β, which captures the average change in a student's test score in the year when they are assigned an HBCU-trained teacher, relative to years when they were assigned a non-HBCU-trained teacher. Because the data are stacked by student-years, with individual students appearing multiple times, and the level of treatment (teacher assignment) varies with and across multiple students, I follow guidance from Cameron, Gelbach, and Miller (2011) on two-way clustering and cluster standard errors at the student and teacher levels.

Second, I also fit a regression using lagged achievement instead of student fixed effects, comparing *across* students assigned to different teachers.

$$A_{it} = \beta(\text{HBCU}_{it}) + \Phi f(A_{it-1}) + \Omega f(A_{-i(t-1)}) + X'_{ijst}\gamma + \theta_g + \eta_t + \epsilon_{it}. \qquad (2)$$

Specifically, I condition on nonlinear (cubic polynomial) functions of both lagged own test scores and lagged leave-out classroom test scores (i.e., the average of all other students' prior test scores) — Φ and Ω are vectors of parameters that correspond with each term of the functions. I also include grade and year fixed effects and additional student-level controls (student gender, age, lagged student absences, lagged suspension receipt, and an indicator for a repeated teacher). The remaining variables and standard error clustering are the same as Equation (1). While lagged achievement regressions are relatively popular in the literature, a drawback in this setting is it necessarily requires non-missing student test scores from the prior year, which is only feasible in 4th and 5th grade (end-of-grade testing begins in 3rd grade in North Carolina), removing a large share of students from the sample.

5 Results

In this section, I present empirical evidence of HBCU-trained teacher positive effects on Black student test scores. The main findings for math test scores, discussed in Section 5.1, are invariant to controls for observable student, teacher, and classroom characteristics, other modeling strategies, and different sample restrictions (Section 5.2). In Section 5.3, I argue the OLS results I present warrant a causal interpretation, following evidence from a quasi-experimental design adapted from Chetty et al. (2014). I demonstrate these teachers help to mitigate the Black-White test score gap (Section 5.4). Turning to the same-race teacher effects literature, I show HBCU-trained teacher effects are independent of racial match, as there are differences in effectiveness between same-race teachers conditional on whether a teacher attended an

HBCU (Section 5.5). Lastly, I extend the analysis to non-testing outcomes by showing HBCU-trained teachers are also associated with a lower probability of suspension for Black students (Section 5.6). Suspension results, too, imply race-independent effects for teachers trained at HBCUs.

5.1 Baseline estimates: HBCU-trained teacher effects on Black student test scores

Table 3. HBCU-trained teacher effects on math and ELA test scores

	(1)	(2)	(3)	(4)	(5)	(6)
HBCU-trained teacher effects, Math	0.030***	0.029***	0.028***	0.032***	0.026**	0.041***
	(0.006)	(0.006)	(0.006)	(0.011)	(0.011)	(0.014)
HBCU-trained teacher effects, ELA	0.010**	0.007	0.007*	0.004	0.007	0.001
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.009)
Student and year FE	X	X	X			
Grade and year FE				X	X	X
Teacher and classroom controls	X	X	X	X	X	X
Endogenous controls			X			
Lagged own and peer test scores				X		X
Lagged own and peer test scores in both subjects					X	
Twice lagged test scores						X
Math						
Control group test score mean	-0.482	-0.482	-0.482	-0.479	-0.470	-0.479
R^2	0.859	0.860	0.862	0.625	0.635	0.658
Observations	457,542	457,542	457,542	298,950	263,253	144,172
ELA						
Control group test score mean	-0.426	-0.426	-0.426	-0.426	-0.427	-0.443
R^2	0.866	0.866	0.867	0.611	0.625	0.651
Observations	474,485	474,485	474,485	307,016	278,423	146,012

The dependent variables are the standardized end-of-grade math and ELA test scores. All regressions include controls for student, teacher, and school characteristics. Baseline model controls for a student's economic disadvantage status, disability status, and English language learner status (all in a given year), teacher race/ethnicity, gender, age, and school-level proportions of student race/ethnicity, gender, students receiving free or reduced-price lunch. Teacher-level and classroom-level controls include a quadratic function of years of experience, median score on NTE/Praxis teacher certification exams, the number of classes a teacher taught during the school year, and indicators for having a graduate degree, having a graduate degree from an HBCU, whether the teacher attended college out of state, and whether their undergraduate college is located in the school district where they teach, whether their college was historically a normal school, class year, and proportion variables for how many students were Black, White, Hispanic, female, economically disadvantaged, disability services eligible, and English language learner status. Endogenous student-level controls in Column 3 include an indicator for receipt of a suspension judgment during the year, number of days absent during the school year, whether the student is designated gifted, the share of students in the classroom who receive a suspension judgment or are designated gifted, and whether the teacher is National Board certified. Lagged achievement regressions (columns 4-6) use the controls from column 2, in addition to controls for student gender, age, whether the student had the teacher in a previous year, and lags for suspensions and days absent. Heteroskedasticity-consistent standard errors clustered at the student and teacher levels in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To begin, I recover baseline results by estimating Equation (1) for all Black students. Regression results appear in Table 3, suppressed to the HBCU attendance variable. Column 1 shows the relationship between having an HBCU-trained teacher and math/ELA scores in the baseline model. Test scores are in standard deviation units: a Black student assigned to an HBCU-trained teacher, on average, scores approximately 0.030 standard deviations (SD) higher on their end-of-grade math exam and 0.010 SD higher on their end-of-grade ELA exam than they score in a different year when assigned to a non-HBCU-trained teacher. The math result in particular is comparable to estimates in prior literature estimating same-race teacher match effects on test scores in North Carolina (0.020 SD; Clotfelter, Ladd, and Vigdor, 2007), and Black teacher-Black student match effects reported in Missouri and Tennessee (0.021 SD; Bartanen and Grissom, 2021), Florida (0.030 SD; Egalite et al., 2015), and North Carolina (0.035 SD; Goldhaber and Hansen, 2010).

HBCU-trained teacher effects hold even after accounting for other observable measures of teacher background and classroom composition. Column 2 extends the baseline model to include controls for teacher and classroom characteristics. For teachers, these comprise their level of experience, graduate degree receipt (in general and specifically from an HBCU), teacher licensure test score, the number of classes taught in the school year, whether the teacher attended college out of state, their college's colocation with the teacher's school district, the college's historical normal school status, and their undergraduate institution's selectivity, as measured by its 2001 rating on Barron's admissions competitiveness index. For classrooms, these include classroom-level aggregates of student race/ethnicity, gender, economic disadvantage, and class size. Collectively, then, the estimated HBCU teacher effect on math scores is not driven by HBCU-trained teachers' relatively higher average teaching experience, any institutional familiarity with North Carolina education curricula that may have arisen from attending a college in-state or in the district where they teach, any general comparative advantage from many HBCUs' histories specializing in teacher education, or the composition of a student's peers. In the ELA equation, the magnitude declines by almost one-third (0.01 to 0.007) and just loses significance at conventional levels ($p = 0.10$).

Including additional observable student, teacher, and classroom controls (column 3) that may impact test scores, such as student suspension receipt, number of days absent from school, gifted status, teacher National Board Certification status, and leave-student-out classroom average for suspensions also have little impact on the

estimated effect. These are post-treatment (i.e., teacher assignment) outcomes, however, and may also be influenced by a student's teacher. To mitigate endogeneity concerns, they are omitted from subsequent regressions.

Finally, I consider the lagged achievement models outlined in Equation (2): the baseline (column 4), an extension with controls for cubic polynomials of lagged student and peer scores in the other subject (e.g., controlling for ELA scores in the math regression) (column 5), and controlling for once- and twice-lagged scores in the same subject (column 6). The estimated math effect remains comparably large and statistically significant at the 1% and 5% levels. The ELA effect, conversely, shrinks further and is imprecisely estimated.

Taking stock of differing results for math and ELA: that I find inconsistent evidence of teacher effects on ELA test scores is relatively common across research in education. Proposed explanations include that a disproportionate share of learning in reading takes place at home rather than in school, so school/teacher effects are muted (Jackson et al., 2014), and that state accountability plans and test designs incentivize different levels of test preparation in the classroom, such that teachers put in varying levels of effort to math compared to ELA (Riehl and Welch, 2022). For parsimony, I focus on math teachers and math test score effects throughout the remainder of the paper using variations on the specification in column 2.[10] [11]

5.2 Robustness checks

Across a battery of alternative specifications, I find consistent evidence of positive HBCU-trained teacher effects on Black student math test scores (Appendix Table 5). Following the discussion in the Data Appendix, bias may arise in the estimating strategy in the presence of differential treatment (i.e., teacher assignment) timing. I follow recent literature on TWFE bias and present two separate specifications. First, I estimate the effect of having an HBCU-trained teacher for the first time in the 5th grade, so that results are identified solely from students who ever have one HBCU-trained teacher and thus are not confounded by possible dosage effects. Second, I

[10] Subsequent analyses for ELA test scores are available upon request to the author.

[11] While the lagged test score model is popular in much of the education literature, I use the student fixed effects specification throughout the remainder of the paper given (1) the data constraints with lagged specifications I discussed in Section 4.1; (2) the point estimates here are comparable between student FE and lagged specifications, and (3) the forecast bias tests I conduct below suggest both specifications produce unbiased estimates of HBCU teacher effects on Black students. As such, I defer to the specification that allows me to keep more data for students (i.e., the student fixed effects regressions).

implement the estimator derived in de Chaisemartin and D'Haultfœuille (2020) that is robust to the differential treatment timing concern. Both coefficients are positive, statistically significant and, in fact, larger than the estimated effect sizes in Table 3, suggesting the TWFE specification may be understating the true impact of HBCU-trained teachers on Black student math exam scores.

The main effects are also robust to various analytic sample restrictions. First, I drop student observations from atypical schooling occurrences: students who appear in multiple classrooms during the school year, transfer schools, and/or are assigned to multiple teachers.[12] Second, to examine whether post-undergraduate experiences confound the "true" HBCU attendance signal (e.g., if any relevant teaching skills are developed post-HBCU undergraduate attendance), I focus on teachers who only hold a bachelor's degree. Third, I check whether effects are driven by colleges in North Carolina by only comparing only out-of-state teachers. Fourth, I restrict the sample to teachers who only attended one institution across any degrees obtained, ensuring that effects are not unknowingly conflating a mix of trainings received by attending multiple schools. Fifth, I estimate effects over the "full information" sample of students for whom I observe all three grades, probing any effects of differential student attrition from the sample and changing student composition, which may bias results if I systematically observe students in atypical years of their schooling. Sixth, I reduce the sample based on the number of Black students in a classroom. A large share of teachers in any given year only teaches one Black student (e.g., across all teacher-year observations, the 25th percentile in the distribution of Black students taught is one, and the median is four). Teaching so few Black students may confound the reliability of these estimates as a signal of teacher effectiveness with Black students (Kim, 2022; Staiger and Rockoff, 2010). To address this, I limit the sample to teachers who teach at least three Black students in a given school year. Finally, I show regression estimates that only includes teachers observed in the sample for three or more years, following caution raised in Koedel and Betts (2011) that using the first couple of years of teacher experience to make inferences about teacher quality. This specification has the added benefit that sorting bias can be mitigated by restricting to teachers with multiple classrooms of students over time.

HBCU effects are also not driven by a few "superstar" institutions that produce an inordinate number of high-quality teachers. I perform several "leave-one-out" regres-

[12]These were previously kept for precision, and prior regressions include controls for each of these less common occurrences (10% of observations in the analytic sample).

sions where I re-estimate my preferred specification and drop one college (Appendix Table 3). I do this 10 times, dropping each of the five largest HBCUs (in the sample) and each of the five largest non-HBCUs. All estimates are comparable to the main effect estimated in Table 3, indicating results are not a function of any one college's contribution to average teacher value-added. That said, I note where slight fluctuations in the point estimates occur. For example, the coefficient becomes larger when UNC Charlotte is dropped (shifting from 0.029 to 0.033). If taken literally, it implies UNC Charlotte supplies some of the relatively stronger non-HBCU-trained teachers, as the gap between HBCU and non-HBCU becomes larger when the college is removed. Because I do not observe the full universe of teachers from these colleges (not all graduates teach in North Carolina after graduation), I would exercise caution in interpreting these coefficients as college-level value-added measures. Future research may find interest in more rigorously exploring college-level heterogeneity effects

5.3 Identifying assumption and causal interpretation

Causal identification in this context assumes conditional random assignment of students to teachers. Following Jackson (2018), formally this means

$$\mathbb{E}[\phi_{it-1} \mid \text{HBCU}_{it}, \chi_{ist}] = \mathbb{E}[\phi_{it-1} \mid \chi_{ist}]. \tag{3}$$

where ϕ_{it-1} is a student i's incoming academic ability (i.e., their stock of ability at the end of the year preceding assignment to an HBCU-trained teacher) and χ_{ist} is a vector that holds observable student, classroom, and school characteristics, including student and year fixed effects. Conditional on χ_{ist}, knowing a student's assignment to an HBCU-trained teacher provides no additional information about a student's incoming academic ability. I argue the assumption holds here, for a couple of reasons. First, following studies that show observational data can provide unbiased predictors of teacher effects when controlling for prior student and peer achievement (e.g., Kane and Staiger, 2008, Kane et al. 2013, Chetty et al., 2014 ,and Bacher-Hicks et al., 2019), I present evidence in columns 4, 5, and 6 of Table 3 that HBCU-trained teacher effects hold in regressions that condition on cubic polynomial functions of prior student and peer (classroom leave-student-out) achievement.

Second, I adopt the quasi-experimental approach developed in Chetty et al. (2014) that evaluates forecast bias in teacher value-added by exploiting quasi-random variation in teacher mobility. In this context, different school-grade-year groups of Black

students vary in their exposure to HBCU-trained teachers, variation induced by plausibly exogenous mobility of teachers across grades and schools.[13] Assuming teacher moves are independent of changes in student or school quality, if HBCU-trained teachers have real predictive effects on Black students and value-added estimates are unbiased, then a cohort's change in average achievement when its exposure to HBCU-trained teachers increases should equal the change predicted by the main model's HBCU effect. Formally, for each year t, I re-estimate the main specification (column 2 from Table 3, which estimates Equation (1)) on all years excluding t to obtain a leave-year-out coefficient $\widehat{\beta}_{-t}$ on the HBCU indicator. Let A^{Black}_{sgt} denote the average test score among Black students in school s, grade g, and year t, and let $\text{Share}^{\text{HBCU}}_{sgt}$ denote the cohort's share of HBCU-trained teachers. I define first differences within school-by-grade lines by

$$\Delta A^{\text{Black}}_{sgt} \equiv A^{\text{Black}}_{sgt} - A^{\text{Black}}_{sg(t-1)}, \qquad \Delta\text{Share}^{\text{HBCU}}_{sgt} \equiv \text{Share}^{\text{HBCU}}_{sgt} - \text{Share}^{\text{HBCU}}_{sg(t-1)}.$$

The predicted change based on the VA model is

$$\Delta\widehat{Y}_{sgt} \;\equiv\; \widehat{\beta}_{-t}\,\Delta\text{Share}^{\text{HBCU}}_{sgt}.$$

I then estimate the forecast-bias regression

$$\Delta A^{\text{Black}}_{sgt} \;=\; \alpha \;+\; b\,\Delta\widehat{Y}_{sgt} \;+\; \eta_{st} \;+\; u_{sg}, \tag{4}$$

where η_{st} are school-by-year fixed effects that absorb school-level shocks common to all grades. Standard errors are clustered at the school-by-grade level, and regressions are weighted by the Black cohort size in (s, g, t) to target population-average effects for Black students (i.e., cohorts with more Black students receive more weight). Under no forecast bias, $H_0 : b = 1$ should hold.[14] I report three versions of Equation (4): (i) the baseline with only $\Delta\widehat{Y}_{sgt}$ and school-by-year fixed effects; (ii) an enriched version adding cohort-level changes in teacher mix and characteristics (experience, age, licensure score, graduate degree, and teacher race/gender shares), which I denote ΔZ_{sgt}; and (iii) a version that also includes grade-by-year fixed effects to absorb grade-specific statewide shocks.

[13] In this 11-year sample, 20.9% of HBCU-trained teachers change schools, and, even among teachers who stay in the same school, 19.5% switch to a different grade.

[14] A nonzero intercept α is not directly informative about bias in b and can reflect secular drifts or differencing noise.

Across different specifications in Appendix Table 4, the point estimates cluster between 1.63 and 2.17, all of which are statistically no different from $b = 1$. These estimates are consistent with at most mild attenuation in the main HBCU effect coefficient: if anything, the main model may slightly understate the HBCU effect. Overall, the evidence indicates no detectable forecast bias and supports a causal interpretation of the estimated HBCU teacher effects for Black students.[15]

5.4 HBCU-trained teachers and the Black-White test score gap

For benchmarking, we can also measure HBCU-trained teacher effects as they relate to the persistent Black-White test score gap (Anderson, 2007; Ladson-Billings, 2006). To do so, I estimate a regression equation that takes the form

$$\begin{aligned} y_{igst} = \alpha \; + \; \beta\,\text{Black}_i \; + \; \gamma\,\text{HBCU}_{jt} \; + \; \delta\big(\text{Black}_i \times \text{HBCU}_{jt}\big) \\ + \; p\big(y_{i,t-1}\big) \; + \; p\big(\bar{y}_{-i,t-1,c}\big) \; + \; \mathbf{X}'_{i,t-1}\boldsymbol{\psi} \; + \; \mathbf{T}'_{jt}\boldsymbol{\lambda} \; + \; \mathbf{C}'_{ct}\boldsymbol{\kappa} \; + \; \mu_{sgt} \; + \; \varepsilon_{igst}. \end{aligned} \tag{5}$$

β reports the conditional average test score difference between Black students and other racial groups. I estimate this equation twice, once with the full student sample and one with only Black and White students, the latter regression providing the Black-White test score gap. Accordingly, δ expresses the change in that gap when students are assigned to an HBCU-trained teacher. With school-by-grade-by-year fixed effects, the estimated changes in the gaps are identified by students in the same grade, school, and academic year who vary only in their assignment to teachers from HBCUs. I display estimates from Equation (5) in Table 4.

[15] As a robustness check, I repeat the exercise using the lagged-score VA model from Equation (2) to form $\widehat{\beta}_{-t}$. The corresponding forecast-bias slopes are very similar: $\widehat{b}$ ranges from 1.56 (baseline) to 2.09 (with cohort controls) to 2.07 (with cohort controls and grade-by-year FE), and in all cases I fail to reject $b = 1$ at conventional levels.

Table 4. Effect of HBCU-trained teacher assignment on racial test score gaps in math

	(1)	(2)
Black students	-0.098***	-0.095***
	(0.002)	(0.002)
HBCU-trained teacher	0.002	-0.006
	(0.008)	(0.009)
HBCU × Black	0.010**	0.019***
	(0.005)	(0.006)
All students	X	
Black and White students		X
Lagged own and peer test scores	X	X
School-by-grade-by-year fixed effects	X	X
Teacher and classroom controls	X	X
R^2	0.73	0.74
Observations	1,245,112	953,066

Notes: The dependent variable is the standardized end-of-grade math test score. Heteroskedasticity-robust standard errors clustered by teacher and school-by-grade-by-year in parentheses. Stars denote significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

HBCU-trained teachers have substantial impacts on the Black-White test score gap. Estimates show the average, within-school Black student vs. all students test score gap (-0.098 SD after controlling for lagged and peer achievement) is reduced by 10% when Black students are assigned to HBCU-trained teachers. The effects are even more dramatic among the Black and White student sample in column 2: HBCU-trained teachers reduce the Black-White test score gap by 20% (0.019/0.095). Further, given the coefficient on the HBCU teacher assignment variable is nearly zero and statistically insignificant in both specifications, these results suggest no negative test score effects on non-Black students assigned to HBCU-trained teachers.[16]

5.5 Heterogeneous effects: Comparing HBCU-trained and non-HBCU-trained teachers of the same race

A straightforward explanation for HBCU effects is that they merely reflect a correlate with teacher race, with results that are driven by Black teachers who also happen to

[16]In related work, Delgado (2025) finds that teachers with a comparative advantage in raising Black students' test scores have no discernible effect on non-Black-students' test scores.

have attended an HBCU. All previous regressions control for teacher race, so empirically this seems unlikely from the start. I more explicitly draw further distinction between these channels and strengthen the evidence of HBCU teacher effects by disaggregating the analysis of HBCU-teacher effectiveness for Black students by teacher race. Empirically, the first approach interacts HBCU attendance and teacher race:

$$\begin{aligned} A_{it} = \alpha + \beta_{WH}\,\text{WhiteHBCU}_{jt} + \beta_{BN}\,\text{BlackNonHBCU}_{jt} + \beta_{BH}\,\text{BlackHBCU}_{jt} \\ + X'_{it}\gamma + Z'_{jt}\delta + C'_{st}\phi + \theta_i + \eta_t + \varepsilon_{it}. \end{aligned} \tag{6}$$

A vector of student-, teacher-, classroom-, and school-level controls (as in column 2 of Table 3) is represented by X. The β coefficients capture the relationships between teacher assignment and student test scores for different types of teachers relative to a White teacher who did not attend an HBCU. Table 5 contains the point estimates for β_{WH}, β_{BH}, and β_{BN}.

Table 5. Heterogeneous effects of HBCU-trained teachers on math test scores, by teacher race

	(1)
Black teacher, HBCU	0.013**
	(0.006)
Black teacher, non-HBCU	-0.012**
	(0.005)
White teacher, HBCU	0.043***
	(0.011)
Student and year FE	X
Teacher and classroom controls	X
Comparing effects	
Black, HBCU vs. Black, non-HBCU	0.025***
	(0.007)
Control group test score mean	-0.471
R^2	0.860
Observations	457,542

The dependent variable is the standardized end-of-grade math test score. Regression includes controls for a student's economic disadvantage status, disability status, and English language learner status (all in a given year), teacher gender, age, school-level proportions of student race/ethnicity, gender, students receiving free or reduced-price lunch, a quadratic function of the teacher's years of experience, median score on NTE/Praxis teacher certification exams, the number of classes a teacher taught during the school year, and indicators for having a graduate degree, having a graduate degree from an HBCU, whether the teacher attended college out of state, whether their undergraduate college is located in the school district where they teach, whether their college was historically a normal school, class size, and classroom proportion variables for how many students were Black, White, Hispanic, female, economically disadvantaged, disability services eligible, and English language learner status. Heteroskedasticity-consistent standard errors clustered at the student and teacher levels in parentheses. *** $p < 0.01$.

Black teacher effectiveness in part varies with HBCU attendance: while Black students in classrooms with Black teachers who attended an HBCU score on average

0.013 standard deviations higher on math exams (compared against White teachers from non-HBCUs), non-HBCU-trained Black teachers perform considerably worse, with students scoring 0.012 standard deviations *lower* than when paired with White, non-HBCU-trained teachers. To make this point more explicitly, I run post-estimation Wald tests to compare β_{BH} and β_{BN} directly. These confirm a large, statistically significant gap in teacher performance between Black teachers from HBCUs and Black teachers from non-HBCUs ($\beta_{BH} - \beta_{BN} = 0.025$, SE = 0.004, $p < 0.001$). In addition to highlighting the within-race effect, note that I only find a positive same-race match effect for Black teachers and students *only* when I condition on that Black teacher having attended an HBCU. That I find no general positive effect for Black teachers with Black students may warrant further investigation in the literature on same-race teacher effects, which tends to find positive effects from same-race teacher pairings, particularly for racial and ethnic minorities (e.g., Egalite, 2023; Redding, 2019). Admittedly, the differences between our results could potentially arise through a number of channels, including differences in geographic settings and time frames, research designs, and outcomes of interest. Even so, I am also not the first to find a null or negative same-race teacher effect for Black students (e.g., Buddin and Zamarro, 2009; Ehrenberg, Goldhaber and Brewer, 1995; Jennings and DiPrete, 2010; Morgan and Hu, 2023). Likewise, when estimating Equation (1) for White, Hispanic, and Asian students paired with teachers of their same race/ethnicity, I only find significant same-race teacher effects for White students.

Next, I draw attention to β_{WH}, which compares HBCU-trained White teachers to White teachers who attended other institutions. Black students in classrooms with White teachers who attended an HBCU score on average 0.043 standard deviations higher on math exams than they do with non-HBCU-trained White teachers. In addition to reinforcing the mounting evidence of the important roles HBCUs play in Black student achievement, these results imply HBCUs have a race-independent effect: for students, the relevant characteristic for improving academic achievement may not simply be a racial match with their teacher, but rather the teachers' prior experiences, including their teacher preparation.

5.6 Beyond test scores: the relationship between HBCU-trained teachers and student suspensions

Positive HBCU-trained teacher effects on Black student achievement may arise from teacher behavior and their classroom management practices. Recent survey data

focused on thousands of teachers of color show that, even among Black teachers, HBCU graduates reported the highest amount of engagement with students (more hours spent on mentoring and counseling and more students allowed to socialize in their classrooms outside regular class time) (Bristol and El-Mekki, 2022). By contrast, suspensions carry well-documented harms – lower test scores, higher dropout risk, and downstream justice involvement – and tend to fall disproportionately on Black students (e.g., Pope & Zuo, 2023; Lee et al., 2011; Bacher-Hicks, Billings & Deming, 2024), who are suspended at roughly three times the rate of White peers (U.S. OCR, 2014). Analyzing this channel is indispensable given my focus on Black students, for whom over one-fourth of students in the sample receive suspensions at some point. I estimate HBCU-trained teacher impacts on Black student suspensions in Equation (6) .

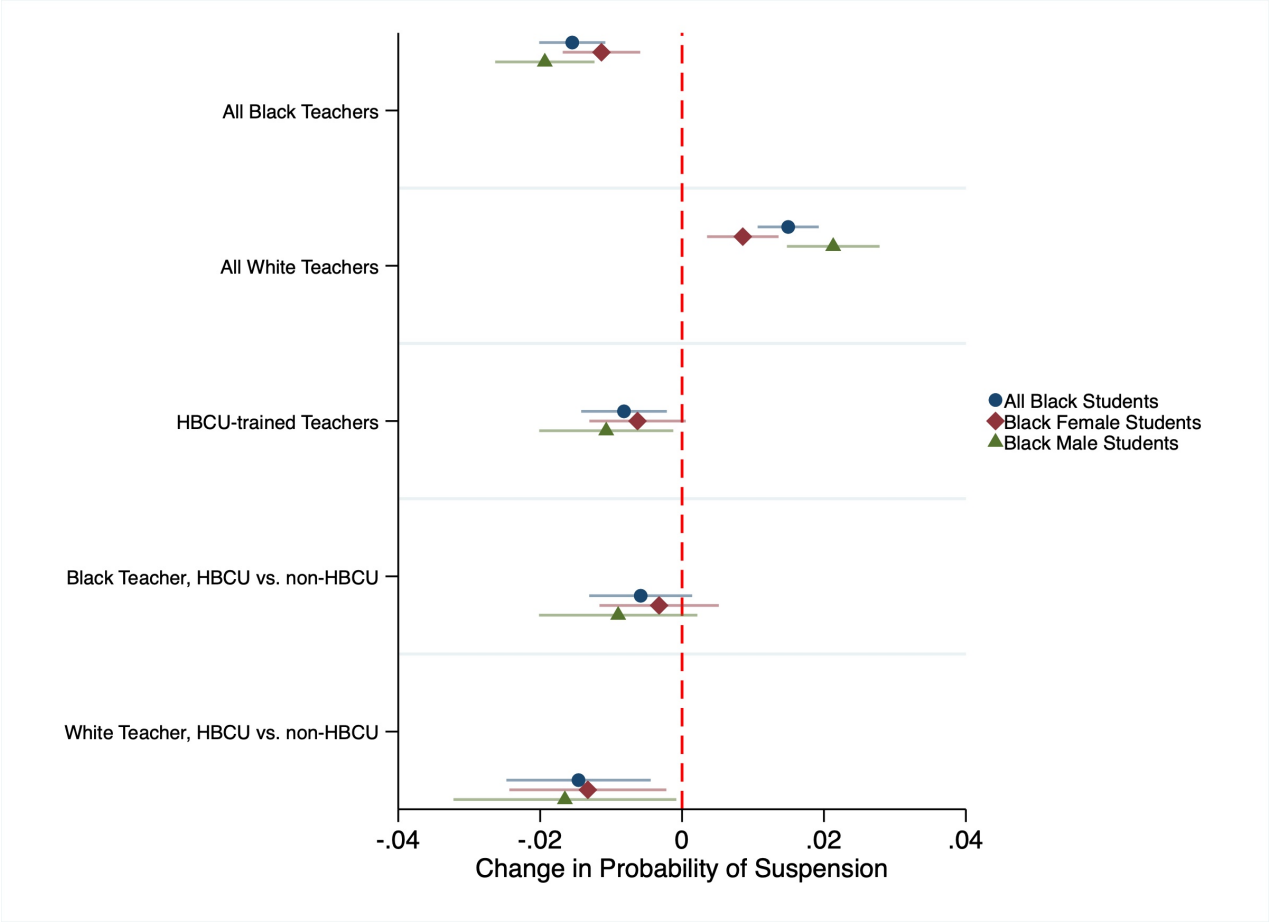
$$\text{Suspension}_{it} = \rho(\text{HBCU}_{it}) + \tau(\text{Suspension}_{it-1}) + W'_{ijct}\phi + \psi_g + \omega_{st} + \zeta_{it}. \quad (7)$$

I use a linear probability model (LPM) to predict the probability that a Black student i is suspended in a given school-year t given variation in teacher assignment.[17] Specifically, I regress whether the student has a suspension judgment in a given year on an indicator for teacher assignment (HBCU-trained or not), an indicator for receiving a suspension in the previous school year, time-varying student-, teacher, and classroom-level covariates, captured in vector W, and grade and school-by-year fixed effects. The school-by-year fixed effects account for any changes in school administration and leadership over time that may influence school suspension rates and place students into comparable contexts.[18] Figure 1 plots coefficients representing the association between HBCU-trained teacher assignment and the probability of being suspended for Black students, with the analogous relationships for Black and White teachers included for reference.

[17]With a relatively short panel — students are typically observed for no more than three years — and thousands of school-by-year and grade fixed effects, a nonlinear model such as a logit or probit would produce inconsistent estimates for the average marginal effect of interest (Angrist and Pischke, 2008; Greene, 2004).

[18]For instance, prior work highlights principals have discretion in disciplinary decisions and overall school discipline (Bacher-Hicks, Billings, and Deming, 2024; Sorenson, Bushway, and Gifford, 2022).

Figure 1: Linear probability model estimates predicting suspension for Black students



The dependent variable is an indicator for whether a student receives a suspension in a school year. Regressions include controls for student gender, economic disadvantage status, disability status, English language learner status, prior year suspension judgment receipt, indicators for a student's quartile ranks of prior year's test score (based on quartile in grade distribution for the entire state in the prior school year) and for a teacher's quartile rank in the number of Black students taught in the prior year, whether a student had the teacher in the previous year, chronic absence in the prior school year (18 or more absences), teacher race, teacher gender, indicators for teacher age, indicators for teacher experience levels (2, 3, 4, and 5+, with 1 year of experience as the omitted category), attending an HBCU for graduate school, teaching multiple classes in a year, large class size (greater than the 90th percentile of class size), classroom demographics (race, gender, economically disadvantaged, disability status eligible, English language learner status), and grade and school-by-year fixed effects. Zero line corresponds to 90% confidence interval. Heteroskedasticity-consistent standard errors clustered at the student and teacher levels.

Consistently, HBCU-trained teacher assignment is associated with a decline in Black student suspension, even after controlling for the teacher's race.[19] When separating results by student gender, I observe larger and more precise effects for boys

[19]Appendix Figure 1 demonstrates a negligible relationship between HBCU teacher assignment and suspensions for White students; and a slightly negative relationship for Hispanic students.

than for girls. Assignment to an HBCU-trained teacher for Black boys is associated with approximately a 4.6% decline in the probability of receiving a suspension (≈ 1.1 percentage point decline from the mean suspension rate of 23.9% for Black boys with non-HBCU-trained teachers).

Similar to the teacher race heterogeneity analysis in Section 5.5, HBCU training is associated with fewer suspensions of Black students even within teacher race, particularly among White teachers. Black boys are 7.1% (1.7 percentage points) less likely to be suspended if they receive a White HBCU-trained teacher over a White teacher from another institution type. Results align with the HBCU training hypothesis: these teachers may have a particular approach to teaching and classroom management that uniquely correlates with reduced disciplinary actions against Black students. According to Irvine (1989), students are disadvantaged when placed with teachers who "do not understand minority student's behavior, physical movements, verbal and nonverbal language, values, worldview, home environment, and learning styles" (pgs. 55-56). In turn, Townsend (2000), in her study of the disproportionate disciplinary actions faced by Black students, recommends teacher training that prioritizes cultural competency as a remedy to address cultural misalignments. Correspondingly, White HBCU-trained teachers may suspend Black students less than other White teachers due to training that mitigates cultural discontinuities and the escalation to discipline.

6 Discussion of mechanisms

6.1 Evidence on selection into HBCUs using survey data

While teacher training theoretically explains the observed HBCU effects, selection into college remains a potential confounder. Given HBCUs often foreground interests in social justice, equity, and community/civic engagement, their students who become teachers may have selected in on positive racial attitudes and/or come from socially progressive backgrounds. If these traits correlate with effective instruction of Black students, these teachers may have been successful independent of whether they attended an HBCU or not. Alternatively, while this may be the case for some students, many others may be drawn to HBCUs for reasons uncorrelated with their potential effectiveness with Black students, such as college costs or a desire to live near their home. Determining whether these results reflect actual effects of HBCUs on their students requires disentangling these competing stories.

To investigate the selection channel, I incorporate a data set on entering college first-years. The Freshman Survey (TFS) is a national, longitudinal (by college) study administered by the Higher Education Research Institute (HERI), dating back to 1966. This large, comprehensive survey collects extensive background and preferences data on a sample of college freshmen. Though the sample of participating colleges is inconsistent over time — not all colleges participate, and participating colleges are not necessarily sampled every year — its long tenure offers data on roughly 1,900 colleges and over 15 million students. Student-level data are publicly available on the HERI website. Colleges in this data set are anonymized, but, crucially, these data include an indicator for HBCU or not.

6.1.1 Definitions and data construction

Most of the information I use comes from a survey question in which students identify how relevant different factors were in their decision to attend their specific college. Students are presented with a list of possible factors and state whether each item was "Very important," "Somewhat important," or "Not important" in their decision. I focus on five major categories that collectively influence students' college choices. These include:

- College distance from home
- College costs
- College reputation
- Students' networks
- Students' pre-college civic and social attitudes

For the last category, I construct an index from TFS survey questions about incoming students' future goals. For this index, which henceforth I will call the Civic and Racial Justice Orientation (CRJO) Index, I draw from students' stated preferences for "Promoting racial understanding," "Influencing the political structure," "Influencing social values," and "Participating in community action programs." Students offer one of four answers: "Essential," "Very important," "Somewhat Important," and "Not important." I translate student responses into ordinal variables with values between 0 and 3, then implement a one-factor maximum likelihood factor analysis, extracting the first principal factor as a continuous measure of students' civic and racial justice

orientation. I standardize the index within survey year to have mean zero and unit variance.[20]

6.1.2 Predicting HBCU attendance from TFS data

To recover the relationships between HBCU attendance and its possible predictors, I estimate the following regression equation:

$$\begin{aligned}\Pr(HBCU_i = 1) = \alpha \; + \; \beta_1 D_{it} \; + \; \beta_2 N_{it} \; + \; \beta_3 (D_{it} \times N_{it}) \; + \; \mathbf{d}_{it}'\boldsymbol{\delta} \; + \; \theta \, C_{it} \\ + \; \mathbf{X}_{it}^{(N)\prime}\boldsymbol{\eta} \; + \; \mathbf{X}_{it}^{(C)\prime}\boldsymbol{\kappa} \; + \; \mathbf{X}_{it}^{(R)\prime}\boldsymbol{\rho} \; + \; \mathbf{B}_{it}'\boldsymbol{\gamma} \; + \; \mu_t \; + \; \varepsilon_{it}, \qquad (8)\end{aligned}$$

where y_{it} indicates HBCU attendance for student i in year t. The indicator D_{it} equals 1 if *at least one* HBCU lies within 25 miles of a student's home zip code. N_{it} is the standardized (within year) count of non-HBCUs within 25 miles of a student's home zip code. The interaction, $D_{it} \times N_{it}$, allows the association between local HBCU availability and HBCU attendance to vary with the nearby supply of non-HBCUs. The vector $\mathbf{d}_{it}$ includes (logged) distance to the nearest HBCU and distance to nearest non-HBCU. C_{it} is the Civic and Racial Justice Orientation (CRJO) index, standardized within survey year. Following insights from the summary statistics, I include variables that measure the salience of various reasons students considered for selecting their chosen college. $\mathbf{X}_{it}^{(N)}$ collects "network" factors (teacher advice, counselor advice, relatives advice); $\mathbf{X}_{it}^{(C)}$ captures the salience of college costs (cost of attendance, expected financial aid); $\mathbf{X}_{it}^{(R)}$ maps to reputation signals (the college's academic reputation, social reputation, and the college's reputation for graduates getting "good jobs" post-graduation). Additionally, I include baseline controls $\mathbf{B}_{it}$ including gender, indicators for mother and father's college attainment (separately), log median income (in 2010 US dollars), and an academic index, which is the average of a student's standardized high school GPA and SAT/ACT scores. μ_t are survey year fixed effects, and heteroskedasticity-robust standard errors ε_{it} are clustered at the home ZIP code level. The considerable differences I observe between Black and White students in the descriptive statistics motivate my splitting the sample by race when estimating which student characteristics predict HBCU attendance.

[20]For more technical details on survey items and variable construction via factor analysis and reliability checks, see the Data Appendix.

Table 6. Predicting HBCU attendance from college freshmen survey data

	Black Freshmen		White Freshmen	
	(1)	(2)	(3)	(4)
Proximity				
$\geq$ 1 HBCU within 25 mi	0.02740*	0.05150**	0.00214***	0.00395***
	(0.014)	(0.025)	(< 0.001)	(0.001)
z(# non-HBCUs within 25 mi)	-0.00123	-0.00222	0.00016***	0.00029***
	(0.003)	(0.008)	(< 0.001)	(< 0.001)
$\geq$ 1 HBCU $\times$ z(# non-HBCUs)	-0.03103**	0.06061**	-0.01260***	-0.02109***
	(0.016)	(0.024)	(0.003)	(0.005)
ln(distance to closest HBCU)	-0.07900***	-0.09103***	-0.00057***	-0.00093***
	(0.004)	(0.007)	(< 0.001)	(< 0.001)
ln(distance to closest non-HBCU)	0.03813***	0.04400***	0.00029***	0.00044**
	(0.004)	(0.008)	(< 0.001)	(< 0.001)
Pre-college student attitudes				
z(CRJO Index)	0.02916***	0.01992***	0.00029***	0.00035
	(0.001)	(0.006)	(< 0.001)	(< 0.001)
College costs				
z(Cost of attendance)	-0.00584***	0.00212	0.00041***	0.00079***
	(0.002)	(0.006)	(< 0.001)	(< 0.001)
z(Financial aid offer)	-0.02829***	-0.02398***	0.00064***	0.00093***
	(0.002)	(0.007)	(< 0.001)	(< 0.001)
College reputation				
z(Academic reputation)	-0.05046***	-0.05980***	-0.00072***	-0.00106***
	(0.002)	(0.006)	(< 0.001)	(< 0.001)
z(Social reputation)	0.03313***	0.04448***	-0.00082***	-0.00100***
	(0.002)	(0.007)	(< 0.001)	(< 0.001)
z(Employment prospects)	0.01137***	0.00836	0.00001	-0.00020
	(0.002)	(0.006)	(< 0.001)	(< 0.001)
Student network				
z(Teacher advice)	-0.01143***	-0.00710	0.00015**	-0.00006
	(0.002)	(0.006)	(< 0.001)	(< 0.001)
z(HS counselor advice)	0.03984***	0.04666***	-0.00014***	-0.00018
	(0.001)	(0.005)	(< 0.001)	(< 0.001)
z(Relatives wanted me to attend)	-0.02426***	-0.03055***	-0.00010	-0.00008
	(0.002)	(0.006)	(< 0.001)	(< 0.001)
Student sample				
All students	X		X	
Prospective teachers		X		X
Partial adjusted R^2 for prediction blocks				
Proximity	0.107		0.009	
Pre-college student attitude	0.004		0.0001	
College costs	0.003		0.0005	
College reputation	0.009		0.001	
Student network	0.009		0.00002	
Year FE	X	X	X	X
DV mean, no HBCUs within 25 mi	0.219	0.259	0.0001	0.0001
Adjusted R^2	0.194	0.209	0.012	0.022
Observations	233,528	9,814	3,166,004	264,889

Notes: "$\geq$ 1 HBCU within 25 mi" is an indicator for whether there is at least one HBCU located within 25 miles of a student's home zip code. "z(# non-HBCUs within 25 mi)" is a standardized count variable denoting the number of non-HBCUs located within 25 miles of a student's home zip code. "$\geq$ 1 HBCU $\times$ z(# non-HBCUs)" is the interaction term between the previous two variables. All regressions use student weights as provided by HERI. All z variables have been standardized to mean 0 and unit variance. Heteroskedasticity-robust standard errors clustered at the home zip code level in parentheses.
Stars denote significance levels: *** (p < 0.01), ** (p < 0.05), * (p < 0.10).

Among the rationales tested, proximity plays an outsized role in predicting HBCU attendance, particularly for White students (Table 6). White students are approximately 0.21 percentage points more likely to select an HBCU when there is at least one HBCU located within 25 miles of their home. This represents about a 20 times increase over the baseline probability of HBCU attendance for White students from areas with no HBCU within 25 miles, though note the very low baseline (P[Attending an HBCU—no HBCU within 25 miles] ≈ 0.0001 among White students). Concurrently, their likelihood is declining in the number of non-HBCUs also located within the same distance. The direction for the HBCU location indicator and the interaction term are the same for Black students. These relationships hold for regressions among all students (columns 1 and 3) and a regression on the sample of surveyed White students who expressed an interest in teaching as their career of choice. For the potential Black teacher sample, the sign on the interaction term actually flips.

To assess the relative importance of the five predictor categories, I also compute the adjusted partial R^2 for each by block of predictors by re-estimating the linear probability models in columns 1 and 3 without each block on the same estimation sample and taking the increase in adjusted R^2 when that block is added back. This penalizes blocks for dimensionality and yields a cleaner comparison across differently sized blocks. Again, proximity stands above the rest as the predominant predictor for HBCU attendance. Removing the proximity block of predictors lowers fit in the specification for White students by $\Delta\bar{R}^2_{\text{Dist}} = 0.009$, accounting for about 75% of the model's explainable variation; distance likewise contributes the most to explained variation in the regression for Black students ($\Delta\bar{R}^2_{\text{Dist}} = 0.107$; $\approx 55\%$).

Note that these results also imply, at best, a minor role for student attitudes in predicting HBCU attendance. For both Black and White students, the CRJO index ranks next to last among the five predictor categories in explaining variation in college choice. Results are then consistent with students selecting into HBCUs for reasons unrelated to factors we may expect predict later teacher effectiveness, consistent with Section 5's results reflecting HBCUs having effects on their would-be teachers.

6.2 Evidence on treatment effects: HBCU effects by college major

If HBCU teacher effects are not driven by selection, do the observed effects reflect treatment? Moreover, are treatment effects *passive* and at the institution level (i.e., affecting all students at the college), or *active*, reflecting specific training received at HBCUs? Using two data sets – (1) HERI data with freshmen from TFS linked to a

senior-year survey and (2) the subset of teachers in the NCERDC data with college major and certification data – I show the documented teacher effectiveness with Black students is downstream of teacher education practices at HBCUs.

6.2.1 Effects on college student attitudes using HERI Senior Survey data

In addition to The Freshman Survey (TFS), HERI administers a follow-up survey, the College Senior Survey (CSS), which dates back to 1994 (and is publicly available through 2008). For a subset of the TFS sample, HERI links students and their responses across the two surveys. Using this, in conjunction with data from the CSS on students' actual college majors, I estimate heterogeneous (by college major) effects on how student sentiment changes over time *within* HBCUs. [21]

$$
\begin{aligned}
z_{it}^{\text{post}} = \alpha \;+\; \beta_1\,\mathbf{1}\{\text{HBCU}_i = 1\} \;+\; \beta_2\, z_{it}^{\text{pre}} \;+\; \beta_3\,\mathbf{1}\{\text{HBCU}_i = 1\}\cdot z_{it}^{\text{pre}} \\
+ \sum_{m\in\mathcal{M}} \Big[\delta_m\,\mathbf{1}\{\text{Major}_{im} = m\} \;+\; \phi_m\,\mathbf{1}\{\text{HBCU}_i = 1\}\cdot\mathbf{1}\{\text{Major}_{im} = m\} \\
+\; \psi_m\, z_{it}^{\text{pre}}\cdot\mathbf{1}\{\text{Major}_{im} = m\} \;+\; \tau_m\,\mathbf{1}\{\text{HBCU}_i = 1\}\cdot z_{it}^{\text{pre}}\cdot\mathbf{1}\{\text{Major}_{im} = m\}\Big] \\
+ \sum_{m\in\mathcal{M}} \pi_m\,\mathbf{1}\{\text{Major}_i^{\text{pre}} = m\} \;+\; \mathbf{X}_{it}'\boldsymbol{\gamma} \;+\; \mu_{\text{CSS}(t)} \;+\; \lambda_{\text{TFS}(t)} \;+\; \varepsilon_{it}.
\end{aligned}
\tag{9}
$$

In Equation (9), z_{it}^{post} is the senior-year CRJO index for student i in year t. The indicator $\mathbf{1}\{\text{HBCU}_i = 1\}$ marks attendance at an HBCU; z_{it}^{pre} is the (centered) freshman CRJO index; $\{\mathsf{Major}_{im}\}_{m\in\mathcal{M}}$ are dummies for realized senior-year major with "Other" as the omitted category; and $\{\mathsf{Major}_{im}^{\text{pre}}\}_{m\in\mathcal{M}}$ are dummies for intended freshman major. Due to small data cells, I group students into five major categories: *education*, *business*, *social sciences*, *biological sciences*, and *humanities*. The vector $\mathbf{X}_{it}$ collect demographic/SES controls (student race/ethnicity, gender, parental education and income, and student's academic index score). CSS year fixed effects $\mu_{\text{CSS}(t)}$ and TFS year fixed effects $\lambda_{\text{TFS}(t)}$ absorb common shocks and survey-wave shifts. Coefficients are interpreted relative to the baseline of a non-HBCU, "Other" major, and $z_{it}^{\text{pre}} = 0$. The term β_1 gives the HBCU–vs–non-HBCU level difference at the baseline; β_2 is the slope of freshman CRJO at non-HBCUs (for "Other"); and β_3 captures how that slope differs at HBCUs (for "Other"). Major main effects δ_m are level shifts for major

[21] With reductions in sample size from linking the surveys, I lack the statistical power to explore trends by race.

m at non-HBCUs, while ϕ_m are *HBCU* $\times$ *major* level shifts—how the HBCU premium/penalty changes for major m. The ψ_m allow the freshman-CRJO slope to vary by major at non-HBCUs, and the triple interactions τ_m indicate how the freshman-CRJO slope at HBCUs differs for that major from its non-HBCU counterpart. I include intended-major indicators π_m, taken from TFS, to mitigate concerns of selection into college majors. Parameters of primary interest for the HBCU $\times$ major heterogeneity are $\{\phi_m, \tau_m\}$. Together with β_1 and β_3 they deliver the HBCU effect for major m at any freshman CRJO level z:

$$\underbrace{\Delta_m(z)}_{\text{HBCU} - \text{non-HBCU for major } m} = (\beta_1 + \phi_m) + (\beta_3 + \tau_m)\, z.$$

Here $(\beta_1+\phi_m)$ is the HBCU–non-HBCU gap for major m among students with average freshman CRJO (since z is centered), and $(\beta_3 + \tau_m)$ shows how that gap changes as freshman CRJO rises.

Table 7. Effect of HBCUs on End-of-College CRJO Index Score by college major

	(1)
HBCU	0.092***
	(0.035)
Freshman CRJO Index	0.765***
	(0.004)
HBCU × Freshman CRJO	0.005
	(0.054)
Major (main effects; baseline: Other Majors)	
Education	0.060***
	(0.011)
Business	-0.082***
	(0.009)
Social Science	0.126***
	(0.012)
Biological Sciences	-0.015
	(0.010)
Humanities	0.071***
	(0.009)
HBCU × Major	
Education	0.141***
	(0.042)
Business	0.029
	(0.078)
Social Science	-0.042
	(0.034)
Biological Sciences	0.012
	(0.043)
Humanities	0.052
	(0.085)
HBCU × Major × Freshman CRJO	
Education	-0.105**
	(0.051)
Business	-0.116*
	(0.060)
Social Science	-0.040
	(0.025)
Biological Sciences	-0.121***
	(0.044)
Humanities	-0.110***
	(0.042)
TFS survey year FE	X
CSS survey year FE	X
R^2	0.535
Observations	177,819

Notes: The dependent variable is the standardized Civic and Racial Justice Orientation (CRJO) index score, constructed from senior year survey responses in the College Senior Survey (CSS). Baseline major is "Other majors." Intended major from TFS and demographic/SES controls included but not shown. Heteroskedasticity-robust SEs clustered by college in parentheses. Stars denote significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

On a high level, estimates point to gains specifically from majoring in education at an HBCU, over studying other fields at these same colleges (Table 7). Beginning with the HBCU main effects, a student at an HBCU who entered college with an average social/racial attitude (CRJO equals zero) and majored in a field outside the five groups listed would experience a shift upwards in their CRJO index score, relative to a comparably racial justice-oriented student in the same group of majors, but at a non-HBCU. Moving down the table, against the reference group of other majors at non-HBCUs, multiple fields appear to have an advantage in raising students CRJO index scores (particularly education, social sciences, and humanities). Most crucially, however, when I interact college major with HBCU attendance, education is the only field with a large, statistically significant increase beyond the general major bump. Education majors at non-HBCUs shift their senior CRJO scores 0.060 SD higher than the outside major at non-HBCUs, while education majors at HBCUs receive a 0.293 SD boost (0.060 + 0.092 + 0.141) relative to the same group. This implies a 0.233 SD increase for HBCU education majors above education majors from other institutions.

6.2.2 Effects on Black student test scores using NCERDC college major data

To link college training directly to Black student test scores, I complement the prior analysis with additional results from the NCERDC data, where over one-third of the teacher records in the analytic sample (38.6%) contain the specific fields undertaken for their undergraduate studies. I define a teacher to hold an education undergraduate major if the teacher's first or second major is in a traditionally educated-related field (e.g., education, early childhood education, school psychology or counseling, etc.). Because I focus on math exam scores, I also flag any teachers who have majors or graduate degrees in Science, Technology, Engineering, or Math (STEM) fields and include controls for whether a teacher received a license in mathematics, elementary education, and/or a temporary teaching license. The primary estimating equation (Equation 10) includes indicators from three of four constructed categories: HBCU graduates who majored in an education field, HBCU graduates who majored in a non-education field, and non-education majors from non-HBCUs (leaving out the non-HBCU education majors as the reference category). $\mathbf{W}_{it}$ collects the time-varying student, teacher, classroom, and school controls used in prior regressions, plus the major and certification indicators.

$$A_{it} = \beta_{\text{HE}}\,\mathbf{1}\{\text{HBCU}_{jt} \times \text{EdMajor}_{jt}\} + \beta_{\text{HN}}\,\mathbf{1}\{\text{HBCU}_{jt} \times \text{non-EdMajor}_{jt}\} + \beta_{\text{NN}}\,\mathbf{1}\{\text{non-HBCU}_{jt} \times \text{non-EdMajor}_{jt}\} + \mathbf{W}'_{it}\gamma + \delta_i + \eta_t + \varepsilon_{it}. \tag{10}$$

Relative to the full sample of teachers (Table 1), the subset of teachers with college major data (Appendix Table 5) preserves the direction of the HBCU to non-HBCU comparisons, although some of the magnitudes differ (e.g., larger share of White HBCU-trained teachers, higher NTE/Praxis exam scores among HBCU-trained teachers). These discrepancies between the two samples motivate my use of entropy balancing, a reweighting method that adjusts the covariate distributions of two groups using constructed weights so that moments of pre-specified characteristics align (Hainmueller, 2012). This allows for better comparability between the full sample and the college major sample. Columns (1) and (2) of Table 8 present the unweighted and entropy-balanced estimates, respectively.

Table 8. HBCU-trained teacher effects on math test scores, by teacher undergraduate major

	(1)	(2)
HBCU-trained teacher, Education major	0.028**	0.032**
	(0.013)	(0.013)
HBCU-trained teacher, non-Education major	-0.009	0.010
	(0.020)	(0.021)
non-HBCU-trained, non-Education major	0.003	0.007
	(0.009)	(0.012)
Student and year FE	X	X
Entropy weights		X
Comparing effects		
HBCU Ed vs. HBCU non-Ed	0.037*	0.023
	(0.021)	(0.021)
HBCU Ed vs. non-HBCU non-Ed	0.025*	0.026
	(0.015)	(0.016)
HBCU non-Ed vs. non-HBCU non-Ed	-0.012	0.003
	(0.020)	(0.022)
R^2	0.881	0.909
Observations	83,754	83,754

Notes: The dependent variable is the standardized end-of-grade math test score. Column (2) applies entropy weights to the specification from column (1). Heteroskedasticity-robust standard errors are clustered at the student and teacher levels. Standard errors for linear contrasts are computed using the delta method, applying the clustered robust variance–covariance matrix from the regression to the specified linear combination of coefficients. Stars denote significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Echoing the findings from the HERI senior survey data, the analysis here reaffirms the importance of education training in HBCUs. HBCU-attending teachers who specifically majored in an education-related program are more effective at raising Black student math test scores than education majors from non-HBCUs. Likewise, within the group of HBCU attendees, education field graduates are more effective than teachers who attended HBCUs but majored in other fields (though admittedly the point estimate is imprecisely estimated in the entropy-balanced specification). This provides another check on selection into HBCUs, as I am essentially conditioning away selection by comparing teachers who made the same college choice (and thus are plausibly comparable on key unobservables).

6.3 HBCUs and effects on Black student achievement: the role of teacher preparation

Why are HBCU-trained teachers more effective with Black students than their non-HBCU attending peers? The large effects even among teachers of the same race imply that HBCU-trained teacher quality is not due to a simple teacher-student racial match. Instead, I propose these teachers draw from *asset-based pedagogy* (ABP), an umbrella term for pedagogical practices that are culturally relevant, responsive, and sustaining (Ladson-Billings, 1995, 2014; Gay, 2002; Paris Alim, 2014; McCarty Lee, 2014; López, 2017; Kwok et al., 2023). Building on decades of qualitative work, quantitative studies find that culturally fluent instruction raises outcomes for minoritized students (e.g., Dee Penner, 2017, 2021; Bonilla, Dee, Penner, 2021). This aligns well with the nature of HBCU teacher education programs, which intentionally embed culturally responsive practices, community engagement, and explicit expectations around socio-emotional learning and equity into training (Arroyo Gasman, 2014; Dilworth, 2012; Toldston and Pearson, 2019; Watkins, 2005). Taken together, these results suggest that HBCU's unique sociocultural components of teacher training advantage their graduates over those from other institutions.

7 Conclusion

This is, to my knowledge, the first quantitative paper studying the effect of HBCU-trained teachers on Black student outcomes. Using various econometric approaches, I show (i) Black elementary school students in North Carolina assigned to HBCU-trained teachers score better on end-of-grade exams and experience fewer suspensions, and (ii) this relationship extends beyond same-race teacher match: both Black and White HBCU-trained teachers are more effective with Black students than their same-race peer teachers from non-HBCUs. Additionally, I offer evidence that the effects likely stem from HBCU-specific training rather than selection into or mere attendance at HBCUs.

Specific to education policy, this work generates at least three major implications. First, that HBCU-trained teacher effectiveness extends beyond racial match should inform policy interventions regarding teacher-student matches. Focus beyond a strict racial match can alleviate both the potential for overburdening Black teachers towards burnout and attrition (Kohli, 2018; Mustafaa, 2023) and concerns among non-Black teachers who may undervalue their capacity for teaching Black students due to their

racial mismatch. Second, this paper demonstrates the need for further study of HBCUs, to better understand the tenets of their teacher preparation programs and how other education schools can integrate relevant practices into their own curricula. Finally, this research highlights the potential benefits of hiring and retaining more teachers trained at Historically Black Colleges, despite the challenges affecting many, including declining enrollment shares (Hinrichs, 2024) and financial distress (Smith-Barrow, 2019). HBCUs are uniquely producing teachers capable of raising Black student outcomes and, consequently, policymakers should grapple with what gains may be lost in the absence of these institutions.

I focus on an educational setting, but these results apply more broadly to other social institutions. The interacting effects of race, training, and disposition on outcomes in this context contribute to our intuition of the gains from same-race pairings. Racial concordance research often focuses heavily on the racial match and implicitly (or explicitly) defers to innateness explanations for observed effects. I propose evidence that the efficacy of racial congruence is less inherent to the paired individual's race, but due to a bundle of skills — in the teacher context, skills like asset-based pedagogy, classroom management practices, etc. — that can be learned and developed, regardless of race. It highlights the need for researchers to be clearer about implicit assumptions and underlying mechanisms when writing and employing models of match effects; otherwise, we can derive incorrect policy solutions. Future work may find value in extending this approach to other settings with an eye towards not *just* racial representation, but also matching on gender and socioeconomic background. Likewise, scholars may push further towards understanding which features are most salient to improving the outcomes of individuals outside of teacher education (e.g., medical patients, legal defendants, etc.), how important skills or traits are acquired, and how that knowledge can best be replicated and disseminated. For instance, Dahl et al. (2021) finds male recruits in a Norwegian military boot camp adopted more egalitarian attitudes about gender after being randomly exposed to female recruits in their squads, but the effects did not persist after treatment. This suggests a role for passive treatment affecting behavior in demographically incongruent settings, while leaving open the possibility that a more active treatment like structured training, as observed in the HBCU context, could impact longer-run outcomes.

References

Abrams, D. S., M. Bertrand, and S. Mullainathan (2012). Do judges vary in their treatment of race? *Journal of Legal Studies 41*, 347–384.

Acosta, M. M., M. Foster, and D. F. Houchen (2018). Why seek the living among the dead? african american pedagogical excellence: Exemplar practice for teacher education. *Journal of Teacher Education 69*(4), 341–353.

Ahn, T., E. Aucejo, and J. James (2021). The importance of matching effects for labor productivity: Evidence from teacher-student interactions. Working Paper 2106, California Polytechnic State University, Department of Economics.

Akbar, R. and M. J. Sims (2008). Surviving katrina and keeping our eyes on the prize: The strength of legacy and tradition in new orleans's hbcu teacher preparation programs. *Urban Education 43*(4), 445–462.

Alsan, M., O. Garrick, and G. Graziani (2019). Does diversity matter for health? experimental evidence from oakland. *American Economic Review 109*, 4077–4111.

Anderson, J. (2007). The historical context for understanding the test score gap. *National Journal of Urban Education and Practice 1*, 1–21.

Angrist, J. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Aronson, J. (2002). Chapter 14 – stereotype threat: Contending and coping with unnerving expectations. In J. Aaronson (Ed.), *Improving Academic Achievement: Impact of Psychological Factors on Education*.

Arroyo, A. T. and M. Gasman (2014). An hbcu-based educational approach for black college student success: Toward a framework with implications for all institutions. *American Journal of Education 121*(1), 57–85.

Ba, B. A., D. Knox, J. Mummolo, and R. Rivera (2021). The role of officer race and gender in police-civilian interactions in chicago. *Science 371*, 696–702.

Bacher-Hicks, A., S. B. Billings, and D. J. Deming (2024). The school to prison pipeline: Long-run impacts of school suspensions on adult crime. *American Economic Journal: Economic Policy 16*(4), 165–193.

Bacher-Hicks, A., M. J. Chin, T. J. Kane, and D. O. Staiger (2019). An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys. *Economics of Education Review 73*, 1–15.

Bacher-Hicks, A. and C. Koedel (2023). Estimation and interpretation of teacher value-added in research applications. In *The Handbook of the Economics of Education*, Volume 6.

Barrett, N., A. McEachin, J. N. Mills, and J. Valant (2020). Disparities and discrimination in student discipline by race and family income. *The Journal of Human Resources 55*(4), 1–51.

Bartanen, B. and J. Grissom (2021). School principal race and the hiring and retention of racially diverse teachers. *Journal of Human Resources*, 1–68.

Bates, M., M. Dinerstein, A. C. Johnston, and I. Sorkin (2023). Teacher labor market policy and the theory of the second best. Working paper.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119*, 249–275.

Bonilla, S., T. S. Dee, and E. K. Penner (2021). Ethnic studies increases longer-run academic engagement and attainment. *Proceedings of the National Academy of Sciences 118*(37), e2026386118.

Borusyak, K., X. Jaravel, and J. Spiess (Forthcoming). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*.

Bristol, T. and S. El-Mekki (2022). Unique impact, unique burdens: Insights into the black male educator experience. Retrieved from `https://www.donorschoose.org/black-male-educator-insights`.

Buddin, R. and G. Zamarro (2009). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics 66*(2), 103–115.

Cherry-McDaniel, M. (2019). Skinfolk ain't always kinfolk: The dangers of assuming and assigning inherent cultural responsiveness to teachers of color. *Educational Studies 55*(2), 241–251.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Clotfelter, C. T., H. F. Ladd, and C. R. Clifton (2023). Racial differences in student access to high-quality teachers. *Education Finance and Policy*, 1–29.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor (2007). How and why do teacher credentials matter for student achievement? Working Paper w12828, NBER.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *The Journal of Human Resources 45*(3), 655–681.

Dahl, G. B., A. Kotsadam, and D.-O. Rooth (2021). Does integration change gender attitudes? the effect of randomly assigning women to traditionally male teams. *The Quarterly Journal of Economics 136*(2), 987–1030.

de Chaisemartin, C. and X. D'Haultfœuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*, 2964–2996.

Dee, T. (2004). Teachers, race and student achievement in a randomized experiment. *The Review of Economics and Statistics 86*(1), 195–210.

Dee, T. (2005). A teacher like me: Does race, ethnicity or gender matter? *American Economic Review 95*(2), 158–165.

Dee, T. and E. Penner (2017). The causal effects of cultural relevance: Evidence from an ethnic studies curriculum. *American Educational Research Journal 54*(1), 127–166.

Dee, T. S. and E. K. Penner (2021). My brother's keeper? the impact of targeted educational supports. *Journal of Policy Analysis and Management 40*(4), 1171–1196.

Delgado, W. (2025). Disparate teacher effects, comparative advantage, and match quality. *Economics of Education Review 106*, 1–17.

Delhommer, S. (2022). High school role models and minority college achievement. *Economics of Education Review 87*.

Dilworth, M. E. (2012). Historically black colleges and universities in teacher education reform. *The Journal of Negro Education 81*(2), 121–135.

Dudley-Marling, C. and K. Lucas (2009). Pathologizing the language and culture of poor children. *Language Arts 86*(5), 362–370.

Edwards, A., J. Ortagus, J. Smith, and A. Smythe (2025). Hbcu enrollment and longer-term outcomes. *American Economic Journal: Economic Policy 17*(3), 197–235.

Egalite, A. and B. Kisida (2018). The effects of teacher match on students' academic perceptions and attitudes. *Educational Evaluation and Policy Analysis 40*(1), 59–81.

Egalite, A. J. (2023). What we know about teacher race and student outcomes. *Education Next 24*(1).

Egalite, A. J., B. Kisida, and M. A. Winters (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review 45*, 44–52.

Ehrenberg, R., D. Goldhaber, and D. Brewer (1995). Do teachers' race, gender, and ethnicity matter? evidence from the national educational longitudinal study of 1988. *Industrial and Labor Relations Review 48*(3), 547–561.

Ferguson, R. F. (2003). Teachers' perceptions and expectations and the black-white test score gap. *Urban Education 38*(4), 460–507.

Gay, G. and K. Kirkland (2003). Developing cultural critical consciousness and self-reflection in preservice teacher education. *Theory Into Practice 42*(3), 181–187.

Gershenson, S., C. M. D. Hart, J. Hyman, N. W. Papageorge, and C. A. Lindsay (2022). The long-run impacts of same race teachers. *American Economic Journal: Economic Policy 14*(4), 300–342.

Gershenson, S., S. B. Holt, and N. W. Papageorge (2016). Who believes in me? the effect of student-teacher demographic match on teacher expectations. *Economics of Education Review 52*, 209–224.

Goldhaber, D., S. Liddle, and R. Theobald (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review 34*, 29–44.

Goldhaber, D., V. Quince, and R. Theobald (2017). Does the match matter? exploring whether student teaching experiences affect teacher effectiveness. *American Educational Research Journal 54*(2), 325–359.

Goldhaber, D., V. Quince, and R. Theobald (2018). Has it always been this way? tracing the evolution of teacher quality gaps in us public schools. *American Educational Research Journal 55*(1), 171–201.

Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the tobit model. *Econometric Reviews 23*(2), 125–147.

Greenwood, B. N., R. R. Hardeman, L. Huang, and A. Sojourner (2020). Physician-patient racial concordance and disparities in birthing mortality for newborns. *Proceedings of the National Academy of Sciences of the United States of America 117*(35), 21194–21200.

Grissom, J. A. and L. R. Keiser (2011). A supervisor like me: Race, representation, and the satisfaction and turnover decision of public sector employees. *Journal of Policy Analysis and Management 30*(3), 557–580.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis 20*(1), 25–46.

Harper, B. and L. Mawhinney (2017). Teacher preparation at historically black colleges and universities: Remaining relevant in a climate of accountability. In L. Mawhinney and E. Petchauer (Eds.), *Teacher Education across Minority-Serving Institutions: Programs, Policies, and Social Justice*, pp. 161–176. New Brunswick, Camden, Newark, New Jersey; London: Rutgers University Press.

Harper, S. R. and I. Davis, Charles H. F. (2012). They (don't) care about education: A counternarrative on black male students' responses to inequitable schooling. *Educational Foundations 26*(1-2), 103–120.

Headley, A. M. and J. E. Wright (2020). Is representation enough? racial disparities in levels of force and arrests by police. *Public Administration Review 80*(6), 1051–1062.

Hill, A. J. and D. B. Jones (2021). Paying for whose performance? teacher incentive pay and the black-white test score gap. *Educational Evaluation and Policy Analysis 43*(3), 445–471.

Hinrichs, P. (2024). An empirical analysis of racial segregation in higher education. *Education Finance and Policy 19*(2), 218–251.

Hoekstra, M. and C. Sloan (2022). Does race matter for police use of force? evidence from 911 calls. *American Economic Review 112*(3), 827–860.

Ingersoll, R., H. May, and G. Collins (2017). Minority teacher recruitment, employment and retention: 1987 to 2013. Learning Policy Institute.

Irvine, J. J. (1989). Beyond role models: An examination of cultural influences on the pedagogical perspectives of black teachers. *Peabody Journal of Education 66*(4), 51–63.

Irvine, J. J. (1990). Black and white students' role models. Unpublished manuscript.

Jackson, K. C. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics 27*(2), 213–256.

Jackson, K. C. (2018). What do test scores miss? the importance of teacher effects on non–test score outcomes. *Journal of Political Economy 126*(5), 2072–2107.

James, J. and J. H. Wyckoff (2022). School segregation, teacher sorting, and the distribution of teachers. Annenberg Institute at Brown University EdWorkingPaper no. 22-659.

Jennings, J. L. and T. A. DiPrete (2010). Teacher effects on social and behavioral skills in early elementary school. *Sociology of Education 83*(2), 135–159.

Kim, E. J. (2022). Signal weighted teacher value-added models. *Statistics and Public Policy 9*(1), 149–162.

Klopfenstein, K. (2005). Beyond test scores: The impact of black teacher role models on rigorous math taking. *Contemporary Economic Policy 23*(3), 416–428.

Koedel, C. and J. R. Betts (2011). Does student sorting invalidate value-added models of teacher effectiveness? an extended analysis of the rothstein critique. *Education Finance and Policy 6*(1), 18–42.

Koedel, C., E. Parsons, M. Podgursky, and M. Ehlert (2015). Teacher preparation programs and teacher quality: Are there real differences across programs? *Education Finance Policy 10*(4), 508–534.

Koefoed, M. S. and E. McGovney (2019). The effect of same-gender or same-race models on occupation choice: Evidence from randomly assigned mentors at west point. *Journal of Human Resources 54*(2), 430–467.

Kohli, R. (2018). Behind school doors: The impact of hostile racial climates on urban teachers of color. *Urban Education 53*(3), 307–333.

Kumar, R. and F. Lauermann (2018). Cultural beliefs and instructional intentions: Do experiences in teacher education institutions matter? *American Educational Research Journal 55*(3), 419–452.

Kwok, A., J. Waddington, J. Davis, S. Halabi, D. Huston, and R. Hemsley (2023). Beginning teachers and strategies for asset-based pedagogy. Technical report. Annenberg Institute at Brown University EdWorkingPaper no. 23-732.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal 32*(3), 465–491.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in u.s. schools. *Educational Researcher 35*(7), 3–12.

Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harvard Educational Review 84*(1), 74–84.

Ladson-Billings, G. and J. D. Anderson (2021). Policy dialogue: Black teachers of the past, present, and future. *History of Education Quarterly 61*(1), 94–102.

Lee, T., D. Cornell, A. Gregory, and X. Fan (2011). High suspension schools and dropout rates for black and white students. *Education and Treatment of Children 34*(2), 167–192.

López, F. A. (2017). Altering the trajectory of the self-fulfilling prophecy: Asset-based pedagogy and classroom dynamics. *Journal of Teacher Education 68*(2), 193–212.

Maylor, U. (2009). 'they do not relate to black people like us': Black teachers as role models for black pupils. *Journal of Education Policy 24*(1), 1–21.

McCarty, T. L. and T. S. Lee (2014). Critical culturally sustaining/revitalizing pedagogy and indigenous education sovereignty. *Harvard Educational Review 84*(1), 101–124.

Mihaly, K., D. McCaffrey, T. R. Sass, and J. R. Lockwood (2013). Where you come from or where you go? distinguishing between school quality and the effectiveness of teacher preparation program graduates. *Education Finance and Policy 8*(4), 459–493.

Morgan, P. L. and E. H. Hu (2023). Fixed effects estimates of student-teacher racial or ethnic matching in u.s. elementary schools. *Early Childhood Research Quarterly 63*, 98–112.

Mustafaa, F. N. (2023). Black educators' racial identity attitudes and culturally relevant pedagogy: A psychological framework and survey of within-race diversity. *American Educational Research Journal 60*(5), 1–35.

Mykerezi, E. and B. F. Mills (2008). The wage earnings impact of historically black colleges and universities. *Southern Economic Journal 75*(1), 173–187.

Nguyen, N., B. Ost, and J. A. Qureshi (2023). Ok boomer: Generational differences in teacher quality. Technical report. Annenberg Institute at Brown University EdWorkingPaper No. 23-831.

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher 41*(3), 93–97.

Paris, D. and H. S. Alim (2014). What are we seeking to sustain through culturally sustaining pedagogy? a loving critique forward. *Harvard Educational Review 84*(1), 85–100.

Paufler, N. A. and A. Amrein-Beardsley (2014). The random assignment of students into elementary classrooms: Implications for value-added analyses and interpretations. *American Educational Research Journal 51*(2), 328–362.

Penney, J. (2017). Racial interaction effects and student achievement. *Education Finance and Policy 12*(4), 447–467.

Pope, N. G. and G. W. Zuo (2023). Suspending suspensions: The education production consequences of school suspension policies.

Price, G. N., W. Spriggs, and O. H. Swinton (2011). The relative returns to graduating from a historically black college/university: Propensity score matching estimates from the national survey of black americans price. *Review of Black Economy 38*(2), 103–130.

Price, J. and J. Wolfers (2010). Racial discrimination among nba referees. *The Quarterly Journal of Economics*, 1859–1887.

Rangel, M. A. and Y. Shi (Forthcoming). First impressions matter: Evidence from elementary school teachers. *Journal of Human Resources*.

Redding, C. (2019). A teacher like me: A review of the effect of student-teacher racial/ethnic matching on teacher perceptions of students and student academic and behavioral outcomes. *Review of Educational Research 89*(4), 499–535.

Roebuck, J. B. and K. Murty (1993). *Historically Black Colleges and Universities: Their Place in American Higher Education*.

Rose, E. K. (2021). Who gets a second chance? effectiveness and equity in supervision of criminal offenders. *The Quarterly Journal of Economics*, 1199–1253.

Sav, G. T. (2010). Funding historically black colleges and universities: Progress toward equality? *Journal of Education Finance 35*(3), 295–307.

Smith-Barrow, D. (2019). Many hbcus are teetering between surviving and thriving. The Hechinger Report, October 23, 2019. Accessed December 30, 2020, `https://hechingerreport.org/many-hbcus-are-teetering-between-surviving-and-thriving/`.

Sorensen, L. C., S. D. Bushway, and E. J. Gifford (2022). Getting tough? the effects of discretionary principal discipline on student outcomes. *Education Finance and Policy 17*(2), 255–284.

Steffensmeier, D. and C. L. Britt (2001). Judges' race and judicial decision making: Do black judges sentence differently? *Social Science Quarterly 82*(4), 749–764.

Toldson, I. A. and D. Pearson (2019). Strategic priorities for historically black colleges and universities with teacher preparation programs. State Higher Education Executive Officers Association, 2019: 1-31.

Townsend, B. L. (2000). The disproportionate discipline of african american learners: Reducing school suspensions and expulsions. *Exceptional Children 66*(3), 381–391.

U.S. Department of Education, Office of Planning, E., P. Policy Development, and P. S. Service (2016). The state of racial diversity in the educator workforce.

U.S. Department of Education, O. f. C. R. (2016). 2013-14 civil rights data collection: A first look. Retrieved from `https://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf`.

Villegas, A. M. and T. F. Lucas (2004). Diversifying the teacher workforce: A retrospective and prospective analysis. *Yearbook of the National Society for the Study of Education 103*(1), 70–104.

Watkins, A. F. (2005). *Cultivating the Education of African American College Students: A Learning Styles Approach*, pp. 122–127. Boston: Pearson Education.

Wilson, V. R. (2007). The effect of attending an hbcu on persistence and graduation outcomes of african-american college students. *The Review of Black Political Economy 34*(1-2), 11–52.

Appendix Table 1: Largest institutions by number of teachers, by college type and race

	All teachers		Black teachers		White teachers	
	School	#	School	#	School	#
Panel A: HBCUs						
1	Fayetteville State University	511	North Carolina Central University	286	Fayetteville State University	204
2	North Carolina Central University	351	Fayetteville State University	283	Elizabeth City State University	137
3	Elizabeth City State University	287	North Carolina A&T State University	260	North Carolina Central University	58
4	North Carolina A&T State University	280	Winston-Salem State University	215	Winston-Salem State University	48
5	Winston-Salem State University	266	Elizabeth City State University	145	North Carolina A&T State University	17
6	Johnson C. Smith University	58	Johnson C. Smith University	57	West Virginia State College	12
7	Shaw University	57	Shaw University	53	Bluefield State College	10
8	St. Augustine's College	52	St. Augustine's College	51	Norfolk State University	5
9	Livingstone College	39	Livingstone College	34	Bowie State University	4
10	South Carolina State University	34	South Carolina State University	30	Livingstone College	4
Panel B: Non-HBCUs						
1	East Carolina University	2,375	East Carolina University	220	Appalachian State University	2,173
2	Appalachian State University	2,248	University of North Carolina at Greensboro	151	East Carolina University	2,111
3	University of North Carolina at Charlotte	1,543	University of North Carolina at Charlotte	131	University of North Carolina at Charlotte	1,364
4	University of North Carolina at Greensboro	1,321	University of North Carolina at Pembroke	64	University of North Carolina at Wilmington	1,169
5	University of North Carolina at Wilmington	1,257	University of North Carolina at Chapel Hill	59	University of North Carolina at Greensboro	1,147
6	Western Carolina University	769	University of North Carolina at Wilmington	46	Western Carolina University	750
7	University of North Carolina at Pembroke	729	Appalachian State University	33	University of North Carolina at Chapel Hill	605
8	University of North Carolina at Chapel Hill	685	North Carolina State University	32	University of North Carolina at Pembroke	461
9	Campbell University	354	North Carolina Wesleyan College	32	Meredith College	335
10	Meredith College	350	University of South Carolina – Columbia	26	Campbell University	333

Appendix Table 2: Degrees in Education and Racial Composition of Graduates

	All colleges		North Carolina colleges		Differences in means			
	HBCU	non-HBCU	HBCU	non-HBCU	(1)-(2)	(3)-(4)	(1)-(3)	(2)-(4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Education share of all graduates								
All students	7.55	7.41	5.87	7.48	0.142	-1.6	1.68**	-0.0654
% Black	6.91	4.68	4.39	4.05	2.23***	0.345	2.52***	0.632*
% White	17	8.77	17	8.21	8.27***	8.79	0.0386	0.552
% Hispanic	8.36	5	10.5	4.11	3.36	6.4	-2.14	0.89
% Asian	2.99	1.91	4.76	1.5	1.09*	3.26	-1.77	0.406
All graduates by race/ethnicity								
% Black	86.5	6.87	83.4	10.7	79.7***	72.7***	3.18	-3.8***
% White	7.88	70.9	11.5	79.3	-63.1***	-67.8***	-3.64	-8.35***
% Hispanic	0.869	6.59	1.22	2.07	-5.72***	-0.85*	-0.354	4.52***
% Asian	0.607	6.61	0.744	3.5	-6***	-2.76***	-0.137	3.11***
Education graduates by race/ethnicity								
% Black	79.2	4.33	62.3	5.77	74.8***	56.5***	16.9	-1.44
% White	17.8	83.9	33.3	87.1	-66.1***	-53.8***	-15.6	-3.2***
% Hispanic	0.962	4.45	2.19	1.14	-3.49***	1.05	-1.22	3.31***
% Asian	0.241	1.7	0.603	0.703	-1.46***	-0.1	-0.363	0.997***
Observations	80	1,483	10	43				

Notes: Institution-level counts are pooled over 2003–2007 before computing shares, ensuring add-up consistency within each school. Reported means are size-weighted by the appropriate denominator (e.g., total graduates, Education graduates, or race-specific totals). Welch unequal-variance two-sample t-tests are used for all differences. Stars denote significance levels: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Appendix Table 3: HBCU-trained teacher effects on Black students, leave one college out

Panel A: HBCUs					
	FSU	NCCU	ECSU	NCAT	WSSU
HBCU-trained teacher	0.025***	0.030***	0.026***	0.026***	0.032***
	(0.007)	(0.007)	(0.006)	(0.007)	(0.006)
Student FE	X	X	X	X	X
Year FE	X	X	X	X	X
R^2	0.861	0.861	0.861	0.861	0.861
Observations	435,168	442,433	442,197	445,114	446,510

Panel B: non-HBCUs					
	ECU	ASU	UNCC	UNCG	UNCW
HBCU-trained teacher	0.028***	0.029***	0.033***	0.028***	0.028***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Student FE	X	X	X	X	X
Year FE	X	X	X	X	X
R^2	0.866	0.861	0.861	0.861	0.862
N	374,919	423,873	421,548	427,527	427,025

Notes: Each specification drops the named college from the estimation sample (leave-one-college-out). Abbreviations: FSU = Fayetteville State University; NCCU = North Carolina Central University; ECSU = Elizabeth City State University; NCAT = North Carolina A&T State University; WSSU = Winston–Salem State University; ECU = East Carolina University; ASU = Appalachian State University; UNCC = University of North Carolina at Charlotte; UNCG = University of North Carolina at Greensboro; UNCW = University of North Carolina at Wilmington. All regressions include student and year fixed effects. Heteroskedasticity-robust standard errors clustered at the student and teacher levels are shown in parentheses. Stars denote significance: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Appendix Table 4: Estimates of forecast bias in HBCU teacher effects

	(1)	(2)	(3)
Predicted test score Δ from HBCU exposure	1.629**	2.173***	2.152***
	(0.722)	(0.823)	(0.757)
School-by-year FE	X	X	X
Grade FE			X
Cohort changes in teacher characteristics		X	X
p-value for F test: $H_0 : b = 1$	0.384	0.154	0.128
Observations	23,195	23,195	23,195

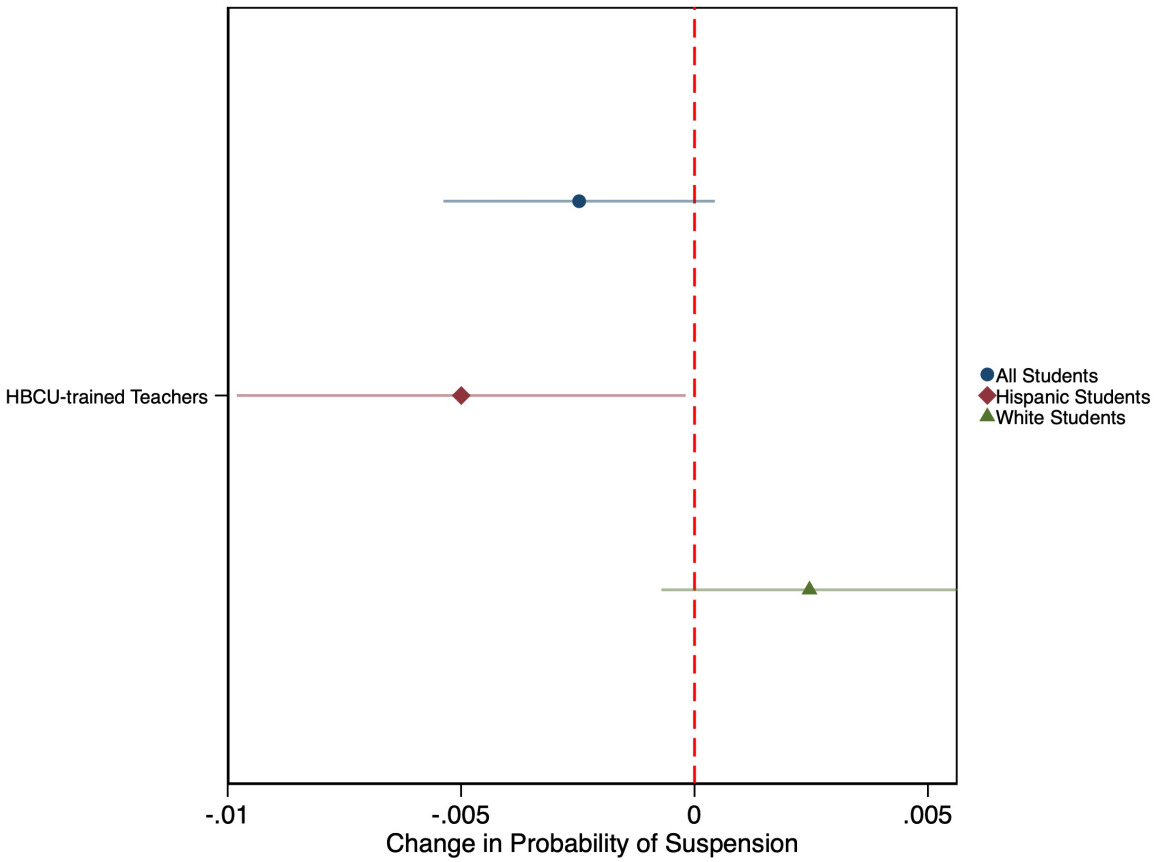
Notes: The dependent variable is the cohort-to-cohort change in mean test score for Black students within school-by-grade grade cells. Columns (2) and (3) include cohort-level changes in teacher characteristics (experience, age, licensure score, graduate degree; and changes in the shares of Black/White/Hispanic/Asian/female teachers). All specifications are weighted by Black-student cohort size. Heteroskedasticity-robust standard errors are clustered at the school-by-grade level. Stars denote significance levels for the $\Delta\widehat{Y}_{sgt}$ coefficient: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Appendix Table 5: Summary statistics for teachers with non-missing college major, 2008–2018

	(1)	(2)	(1)-(2)
Panel A: All teachers			
	HBCU-trained	Non-HBCU-trained	Difference
Black	66.31	7.72	58.59***
White	29.25	87.89	-58.64***
Hispanic	2.30	1.54	0.76
Asian	0.46	0.75	-0.29
Female	90.35	92.11	-1.76
Mean age	35.55	32.47	3.09***
Mean years of experience	5.89	5.44	0.45**
Graduate degree	33.84	31.98	1.86
Graduate degree from an HBCU	8.58	1.24	7.33***
Mean NTE/Praxis z score	-0.27	0.25	-0.52***
Attended college out of state	9.19	26.00	-16.81***
National Board Certified	2.76	4.96	-2.20***
Mean classes taught in year	1.27	1.23	0.05**
Mean number of Black students taught in a year	13.63	8.37	5.26***
Mean share of Black students in school	48.55	29.82	18.73***
Degree in STEM field	3.06	1.67	1.39**
Certification in math	3.37	4.74	-1.37*
Certification in elementary education	98.77	99.52	-0.75*
Temporary certification	7.81	4.62	3.19***
Panel B: Education majors			
	HBCU-trained	Non-HBCU-trained	Difference
Black	60.08	6.47	53.61***
White	34.73	89.50	-54.77***
Hispanic	2.79	1.43	1.37*
Asian	0.60	0.63	-0.04
Female	91.02	92.94	-1.92
Mean age	34.92	31.44	3.48***
Mean years of experience	5.61	5.29	0.32*
Graduate degree	31.74	27.62	4.12*
Graduate degree from an HBCU	7.78	1.14	6.64***
Mean NTE/Praxis z score	-0.14	0.25	-0.39***
Attended college out of state	5.79	21.07	-15.28***
National Board Certified	2.79	5.01	-2.22***
Mean classes taught in year	1.26	1.22	0.04
Mean number of Black students taught in a year	13.03	8.22	4.81***
Mean share of Black students in school	47.11	29.27	17.84***
Degree in STEM field	1.20	0.46	0.74
Certification in math	2.99	4.33	-1.34*
Certification in elementary education	99.60	99.87	-0.27
Temporary certification	1.80	1.51	0.29
Panel C: HBCU-trained teachers			
	Education majors	Non-education majors	Difference
Black	60.08	86.84	-26.76***
White	34.73	11.18	23.55***
Hispanic	2.79	0.66	2.14**
Asian	0.60	0.00	0.60*
Female	91.02	88.16	2.86
Mean age	34.92	37.63	-2.71***
Mean years of experience	5.61	6.84	-1.23**
Graduate degree	31.74	40.79	-9.05**
Graduate degree from an HBCU	7.78	11.18	-3.40
Mean NTE/Praxis z score	-0.14	-0.70	0.56***
Attended college out of state	5.79	20.39	-14.61***
National Board Certified	2.79	2.63	0.16.
Mean classes taught in year	1.26	1.32	-0.06
Mean number of Black students taught in a year	13.03	15.62	-2.60***
Mean share of Black students in school	47.11	53.30	-6.19***
Degree in STEM field	1.20	9.21	-8.01***
Certification in math	2.99	4.61	-1.61
Certification in elementary education	99.60	96.05	3.55**
Temporary certification	1.80	27.63	-25.84***
Observations (Panel A)	653	7,723	
Observations (Panel B)	501	6,304	
Observations (Panel C)	501	152	

Notes: Sample is restricted to teachers in the estimation sample with non-missing college major information. Time-invariant characteristics are constructed at the teacher level within this sample (binaries via max; continuous measures via mean across observed years). Time-varying classroom measures are averaged across years per teacher. Welch unequal-variance two-sample t-tests are used for all differences. Shares (race, female, graduate degree, graduate degree from an HBCU, attended college out of state, National Board Certified, STEM degree, math certification, elementary certification, temporary certification) are displayed in percentage points. Panel B restricts to teachers whose undergraduate major is in education; Panel C restricts to HBCU-trained teachers and compares education majors to non-education majors. Stars denote significance:
*** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Appendix Figure 1: Linear probability model estimates predicting suspension for all, Hispanic, and White students



The dependent variable is an indicator for whether a student receives a suspension in a school year. Regressions include controls for student gender, economic disadvantage status, disability status, English language learner status, prior year suspension judgment receipt, indicators for a student's quartile ranks of prior year's test score (based on quartile in grade distribution for the entire state in the prior school year) and for a teacher's quartile rank in the number of Black students taught in the prior year, whether a student had the teacher in the previous year, chronic absence in the prior school year (18 or more absences), teacher race, teacher gender, indicators for teacher age, indicators for teacher experience levels (2, 3, 4, and 5+, with 1 year of experience as the omitted category), attending an HBCU for graduate school, teaching multiple classes in a year, large class size (greater than the 90th percentile of class size), classroom demographics (race, gender, economically disadvantaged, disability status eligible, English language learner status), and grade and school-by-year fixed effects. Zero line corresponds to 90% confidence interval. Heteroskedasticity-consistent standard errors clustered at the student and teacher levels.

Appendix Table 6: Robustness checks for HBCU-trained teacher effects on math test scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
de Chaisemartin & D'Haultfoeuille (2020b) estimator	0.084***	0.084***								
	(0.028)	(0.013)								
Only students with HBCU-trained teacher in 5th grade			0.036***							
			(0.010)							
Students with one teacher, one classroom, and one school in a given year				0.029***						
				(0.008)						
Teachers with only a bachelor's degree					0.021**					
					(0.009)					
Only teachers who attended college out-of-state						0.041*				
						(0.022)				
Only teachers who attended one institution							0.021**			
							(0.008)			
Students observed in all grades								0.036***		
								(0.008)		
Teachers who teach at least three Black students in a year									0.030***	
									(0.006)	
Teachers observed at least three years										0.032***
										(0.007)
Student FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Control group test score mean			-0.469	-0.461	-0.500	-0.437	-0.497	-0.471	-0.495	-0.449
R^2			0.863	0.859	0.871	0.888	0.868	0.835	0.860	0.865
Observations			351,741	398,143	210,653	60,058	268,291	238,584	430,813	333,429

All regressions include controls for a student's economic disadvantage status, disability status, and English language learner status (all in a given year), teacher gender, age, school-level proportions of student race/ethnicity, gender, students receiving free or reduced-price lunch, a quadratic function of the teacher's years of experience, median score on NTE/Praxis teacher certification exams, the number of classes a teacher taught during the school year, and indicators for having a graduate degree, having a graduate degree from an HBCU, whether the teacher attended college out of state, whether their undergraduate college is located in the school district where they teach, whether their college was historically a normal school, class size, and classroom proportion variables for how many students were Black, White, Hispanic, female, economically disadvantaged, disability services eligible, and English language learner status. Any variables that become collinear due to a given specification (e.g., teacher having a graduate degree for the regression that restricts to only teachers with undergraduate degrees) are necessarily removed. Given the computation of the de Chaisemartin and D'Haultfoeuille (2020b) estimator, test score mean and R^2 are unknown and omitted from this table. This estimator's standard errors were computed using 1,000 bootstrap replications. Rather than cluster at the teacher and student levels for the de Chaisemartin and D'Haultfoeuille estimator, Column 1 clusters at the teacher level and Column 2 results cluster at the student level. For all other specifications, heteroskedasticity-consistent standard errors clustered at the teacher and student levels in parentheses. Stars denote significance: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).

Appendix Table 7: Robustness checks for HBCU-trained teacher effects on math test scores, within teacher race

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Black teacher, HBCU	0.022**	0.012*	0.005	0.052**	0.004	0.014*	0.013**	0.014*	0.026	0.017	0.049
	(0.010)	(0.007)	(0.009)	(0.025)	(0.008)	(0.008)	(0.006)	(0.007)	(0.024)	(0.022)	(0.034)
Black teacher, non-HBCU	-0.014**	-0.011**	-0.006	0.002	-0.008	-0.018***	-0.011**	-0.014**	0.003	-0.002	0.011
	(0.006)	(0.006)	(0.008)	(0.012)	(0.007)	(0.007)	(0.005)	(0.006)	(0.023)	(0.022)	(0.034)
White teacher, HBCU	0.040*	0.044***	0.042***	-0.003	0.037***	0.050***	0.047***	0.046***	0.047**	0.038*	0.050**
	(0.022)	(0.012)	(0.014)	(0.042)	(0.013)	(0.015)	(0.011)	(0.013)	(0.019)	(0.019)	(0.025)
Student and year FE	X	X	X	X	X	X	X	X			
Grade and year FE									X	X	X
Teacher and classroom controls	X	X	X	X	X	X	X	X	X	X	X
Comparing effects											
Black, HBCU vs. Black, non-HBCU	0.037***	0.024***	0.010	0.050**	0.013	0.032***	0.025***	0.028***	0.023*	0.019	0.038**
	(0.010)	(0.008)	(0.011)	(0.024)	(0.010)	(0.009)	(0.007)	(0.008)	(0.013)	(0.013)	(0.016)
Control group test score mean	-0.459	-0.451	-0.493	-0.425	-0.487	-0.461	-0.485	-0.440	-0.468	-0.458	-0.469
R^2	0.863	0.859	0.871	0.888	0.868	0.835	0.860	0.865	0.625	0.635	0.658
Observations	351,741	398,143	210,653	60,058	268,291	238,584	430,813	333,429	298,950	263,236	144,172

The dependent variable is the standardized end-of-grade math test score. Each column reports results from a separate regression of test scores on indicators for having an HBCU-trained Black teacher, a non-HBCU-trained Black teacher, and an HBCU-trained White teacher, restricting the sample to Black students. All specifications include the same set of teacher and classroom controls as in the main specification. Columns 1–8 include student and year fixed effects and correspond, respectively, to the following sample restrictions: (1) students with an HBCU-trained teacher only in fifth grade; (2) students with one teacher, classroom, and school in a given year; (3) teachers with only a bachelor's degree; (4) teachers who attended college out-of-state; (5) teachers who attended one institution; (6) students observed in all grades; (7) teachers who teach at least three Black students in a year; and (8) teachers observed at least three years. Columns 9–11 include grade and year fixed effects and add lagged achievement controls: (9) lagged own and peer math scores; (10) lagged own and peer scores in math and ELA; and (11) twice lagged math scores. The control group test score mean is computed among students taught by non-HBCU-trained White teachers in each specification.
Heteroskedasticity-consistent standard errors clustered at the student and teacher levels are reported in parentheses.
Stars denote significance:
*** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.10$).