This is the final accepted unformatted version of Harbatkin, E. (2021). Does Student-Teacher Race Match Affect Course Grades? *Economics of Education Review*. 81, 102081. https://doi.org/10.1016/j.econedurev.2021.102081.

Does Student-Teacher Race Match Affect Course Grades?

Erica Harbatkin

Michigan State University

Abstract: A growing body of research has found that student-teacher race match is associated with higher test scores, teacher expectations, and teacher perceptions of students. This paper contributes to the student-teacher race match literature by investigating the effect of race match on course grades. To the extent that race match is associated with higher course grades for minority students, a more diverse teacher workforce is one mechanism that may help to narrow the achievement gap. Using a series of fixed effects models exploiting within-student variation across year and subject matter, I find that having a race-matched teacher is associated with a small but significant increase in course grade, on average. The positive effect of race match is driven largely by the experience of Black students.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305E150017 to Vanderbilt University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Does Student-Teacher Race Match Affect Course Grades?

This paper examines the effect of student-teacher race match on the course grades that teachers assign to students. Prior research on student-teacher race match has found Black students tend to be assessed more favorably by race-congruent teachers (Ehrenberg et al., 1995; McGrady & Reynolds, 2013; Ouazad, 2014) and that White math and science teachers assess Black students less favorably than their White students (Takei & Shouse, 2008). These perceptions begin with expectations, with non-Black teachers reporting lower expectations for Black students than Black teachers of those same students (Gershenson et al., 2016). These past studies have focused on teachers' subjective assessments of students using teacher survey measures to operationalize teacher perceptions. Negative teacher perceptions may diminish student outcomes such as exam scores, course taking, and high school graduation (Dee, 2004, 2005; Oates, 2003) because students may calibrate their performance to meet teacher expectations, with teacher expectations thus becoming self-fulfilling prophecies (Brattesani et al., 1984; Farkas et al., 1990; Jussim, 1989; Jussim & Harber, 2005). Oates (2003) found evidence that teacher perception effects may be even stronger for Black students than for White students. Further, teacher perceptions may impact student GPAs, college recommendations, and behavioral incidents—and ultimately college acceptances as a result. If teachers assess different-race students less favorably, the negative consequences will disproportionately affect minority students because most public school teachers are White; students of color account for about 38 percent of the public school student population, but only about 15 percent of public school teachers (Loeb & Reininger, 2004). To that end, the racial composition of the teacher workforce may be an important mechanism for narrowing the achievement gap between minority and White students.

The present study contributes to the body of research on the effects of student-teacher race match on teacher perceptions in three key ways. First, where earlier papers have focused on survey measures of teacher perceptions, I investigate the effects of race match on actual course grades, reducing the social desirability bias that may attenuate effects of survey data as the dependent variable (Bertrand & Mullainathan, 2001; Bound et al., 2001). Second, my analysis expands the literature to fifth through 10th grades, where existing research has focused only on K-5 and high school. Finally, I advance the conversation about student-teacher race match further beyond Black and White by exploring race match effects for Hispanic, Asian, and American Indian students. Additionally, it is worth noting that while a large body of research has examined the effect of student-teacher race match on test scores, this study is the first as far as I am aware to examine the role of student-teacher race match in course grades. Specifically, I ask:

1. To what extent does teacher/student race match affect teacher assessment of student knowledge as measured by course grades?
2. To what extent is this effect moderated by student race?

Drawing from rich statewide administrative data and using quasi-experimental methods, I find that having a race-matched teacher is associated with a small but significant increase in course grade on average, and that the increase is driven largely by the experience of Black students.

Literature Review

Existing research on student-teacher race match has focused largely on student outcomes such as test scores (Dee, 2004, 2005; Jussim, 1989), disciplinary incidents (Bates & Glick, 2013; Downey & Pribesh, 2004; Rocque, 2010), and teacher perceptions and expectations as measured by teacher surveys (Bates & Glick, 2013; Dee, 2005; Downey & Pribesh, 2004; Gershenson et al., 2016; Jussim et al., 1996; McGrady & Reynolds, 2013; Oates, 2003; Takei & Shouse, 2008). This research has largely—though not entirely—found that minority students with race-matched teachers fare better than their peers with different-race teachers. Test score effects have been small—approximately .01 to .015 standard deviations (Egalite et al., 2015) or 4 to 5 percentile points (Dee, 2004, 2005). Black students have experienced the largest positive effects on test scores (Joshi et al., 2018).

The literature suggests two ways in which student race and ethnicity can impact student experiences and outcomes. First, biases exerted by teachers themselves, whether conscious or unconscious and whether prejudiced for or against members of their own race, have been termed "active teacher effects" in the literature (Dee, 2005; Quillian, 2008; van Ewijk, 2011). One form of these active teacher effects could be in-group bias—or endophilia—where teachers would favor students who share their own characteristics (Feld et al., 2016; Jansson & Tyrefors, 2018; Lavy et al., 2018). Second, role model effects in which having a demographically similar teacher increases student motivation and stereotype threat in which students adjust their effort or performance because of a stereotype about their racial or demographic group that may be made salient by the teacher's (shared or unshared) race have been termed "passive teacher effects" in the literature (Adair, 1984; Dee, 2005; Evans, 1992; Foster, 1993; Hess & Leal, 1997; Steele & Aronson, 1995). Each of these potential mechanisms may apply to characteristics such as race and ethnicity (e.g., Bates & Glick, 2013; Dee, 2004, 2005; Egalite et al., 2015; Gershenson et al., 2016; Ouazad, 2014), gender (e.g., Dee, 2005; Feld et al., 2016; Jansson & Tyrefors, 2018), socioeconomic status (e.g., Kozlowski, 2015; Vinopal, 2020), or religion (e.g., Lavy et al., 2018).

The conceptual framework for the present study draws from an unbiasedness benchmark outlined by Ferguson (2003), which requires that teacher expectations are based on observable predictors of performance, such as previous grades and test scores. Similar to other studies on student-teacher race match, my empirical approach works from this benchmark—first by including a vector of rich covariates, including exam scores, that may affect student course grades, and then estimating student fixed effects models to focus on within-student variation and thereby move beyond the requirement that predictors of performance be observable.

Any of these mechanisms might drive the effect estimates in this analysis. For example, if White teachers assign lower course grades to Black students due to unconscious bias while Black teachers rate those students more favorably, the race match effect would be positive for Black students. However, if White teachers have lower expectations for Black students and therefore award them higher course grades for lower levels of mastery than their White peers, the race match effect could be negative for Black students. These active teacher effects could therefore operate in either direction. Passive teacher effects could occur if, for example, Black students exert more effort for Black teachers than for White teachers, contributing again to a positive race match effect for Black students. In-group bias could contribute to a positive race match effect if teachers make a stronger effort to support same-race students and reward them with higher course grades. Controlling for exam score will account for any part of these mechanisms that influences both course grade and exam performance.

Data, Sample, & Measures

This analysis uses North Carolina administrative data over three school years from 2014-15 through 2016-17. The sample is middle and high school students across all North Carolina public schools (5th through 10th grade) taking the state accountability exams for math, reading, or science, who have course grades for those subjects. I link exam scores to course-level data based on course code and a unique student identifier. I then merge in student demographic data, teacher demographics and experience, and school-level variables.

Sample

The analytic sample includes student-course combinations in which the student has a course grade and accountability exam score. The data contain 3.8 million student-course observations with a course grade and accountability exam score. I drop 59,314 observations in which students have more than one record with different course grades for the same course. These students are likely taking a double dose, or two course periods, of the tested course (Henry et al., 2016). I then drop 18,647 course records for students who have multiple teachers for a given course because the analysis incorporates teacher-level covariates. A small number of students are coded in different race categories across multiple observations. In these cases, if a student has a value of "Two or more races," or missing then I reassign all records for that student to the specified nonmissing race. I then drop the remaining 2,502 students who are coded with different race categories. Finally, I exclude 59,016 students and 2,672 teachers[1] whose race is missing or categorized as "Two or more races" (both teachers and students) or "Other" (teachers only) because I cannot generate a credible race match value for them.[2] The final analytic dataset has about 3.2 million student-course observations, 876,534 unique student records, 25,465 unique teachers, and 2,082 schools.

[1] These 2,672 teachers have 455,639 students assigned to them.

[2] For example, a teacher coded as "Two or more races" might be Black and Asian and therefore race match with both Black and Asian students, but coding "Two or more races" as its own category would incorrectly classify that teacher's Black and Asian students as not race-matched. Meanwhile, because teachers can be coded as "Other" while students cannot and because I cannot observe the race of a teacher coded as "Other," dropping these teachers reduces the chance of a similar coding discrepancy.

Measures

Dependent variable. The dependent variable for this analysis is the student's A-F course grade in the tested subject. Because some course grades are recorded in the state dataset as letter grades and others as numbers from 0 to 100, I convert all numeric grades to letters, with 90 and above as *A*, 80–89.9 as *B*, 70–79.9 as *C*, 60–69.9 as *D*, and below 60 as *F*. I exclude grades of pass, incomplete, or withdraw. Letter grades are then converted to a GPA scale, with an *A*'s as 4 points, *B*'s as 3, *C*'s as 2, *D*'s as 1, and *F*'s as zero points.[3] About 70 percent of grades assigned were *A* or *B* and about 15 percent were *D* or *F*. The mean grade on the GPA scale is about 2.73.

Table 1 ABOUT HERE

Teachers are responsible for assigning course grades for their students. In courses with EOCs (i.e., Integrated Math 1, Biology, and English 2), state policy requires that the EOC assessment count toward at least 20 percent of a student's grade. Because of this requirement, I expect that any race match effect in EOC courses will be attenuated. Additionally, the letter grade outcome is a coarser measure than a continuous course grade and I therefore expect estimates using this strategy to be less precise. In robustness checks, I estimate on two alternative course grade operationalizations. I first run models estimating a set of dichotomized letter grade variables reflecting whether the grade is an *A*, *B* or above, *C* or above, and *D* or above. I then predict the continuous numeric grade for just the 2.7 million student-course records representing 805,838 students in the subset of 1,985 schools that report numeric grades. The mean numeric course grade for these records is 81.4 on a 100-point scale and has a standard deviation of 13.1.

[3] I take the four-point approach to align with calculation of grade-point averages, which are a heavily weighted factor in many college applications, and therefore have broad policy relevance. While I would ideally use grade points that account for pluses and minuses (e.g., A-, B+, B-, etc.), I am unable to do so because district policies around pluses and minuses are not consistent across the state, with some districts either not assigning plus and minus grades or not entering pluses and minuses in the dataset.

Student-teacher race match. The independent variable of interest for the first research question is an indicator for whether the student and teacher are the same race. For the second research question, which aims to examine heterogeneous effects of race match by student race, the independent variable of interest is a linear combination of the *race match* indicator and the student race. As shown in the first column of Table 2, bottom row, students and teachers were race-matched in more than half of the student-course observations, though these proportions vary substantially by student race. White students were race-matched most often at more than 90 percent of the time. American Indian students were race-matched in about one-third of their classes, Black students about one-fourth, and Hispanic and Asian students 1 and 2 percent, respectively. These figures indicate high levels of racial segregation of both students and teachers. School demographics show that White teachers tend to teach in schools with majority-White student populations, while Black, Hispanic, Asian, and American Indian teachers work in majority-minority school settings (Appendix Table A-1 provides school-level student demographics by teacher race).

Table 2 ABOUT HERE

The second two columns of Table 2 show that students with race-matched teachers receive a mean numerical course grade of about 2.89 while their counterparts without race-matched teachers have a mean course grade of just 2.52. Descriptively, White students who have race-matched teachers receive higher course grades than their counterparts without race-matched teachers, while Black students who have race-matched teachers receive *lower* course grades than their counterparts without race-matched teachers. But Black students are more likely to be race-matched in lower performing schools, which biases the overall relationship between race congruence and course grade for Black students downward. White students, meanwhile, are more likely to be race-matched in higher performing schools and less likely to be race-matched in lower performing schools, where there are more teachers of color.[4] These patterns in the raw data, which run counter to the race match effects in the literature, may be explained by a number of factors, including residential segregation leading Black students to be race matched more often in underresourced schools and White students to be race matched more often in affluent schools.

Covariates. All models include the student's score on the state accountability exam linked to the course subject. North Carolina requires districts to administer exams in math and reading for third through eighth graders, and in science for fifth and eighth graders. These third- through eighth-grade exams, which are called end-of-grade exams (EOGs), align with state content standards for the subject and grade. The state requires districts to administer three high school accountability exams—Integrated Math 1, English 2, and Biology. These exams, called end-of-course exams (EOCs), align with subject-specific content standards. I link exam scores to course grades based on the exam associated with that particular course. Both EOGs and EOCs are administered in the last 10 days of the school year.[5] State policy calls for school districts and vendors to score exams. Teachers do not score their students' exams.[6] The state Board of Education requires that EOCs count 20 percent toward a student's final course grade, while there is no state requirement for EOGs to count toward a student's final grade.

---

[4] The simple correlation between the race match indicator and the school performance composite for Black students is -0.25, indicating that as school performance increases Black students are less likely to have race matched teachers. The correlation goes in the opposite direction for white students (.13).

[5] For EOC subjects that are taken during a single semester (e.g., half-year block classes), the testing window is the last five days of the semester.

[6] All EOG and EOC exams were developed, piloted, and validated by the state to align with state content standards. All test items undergo a review process to ensure they align with content standards. EOG test items are multiple choice (reading, math, and science) and numeric entry (math). Test items on the Math 1 and Biology EOCs are all multiple choice, while the English 2 EOC contains multiple choice and constructed response questions. State requirements call for school districts to score exams with only multiple choice and numeric entry items, and for test vendors to score exams that include constructed response items.

I standardize exam scores by subject and grade level to have a mean of 0 and standard deviation of 1. By controlling for exam score, I am able to condition each student's course grade on her or his mastery of the subject as measured by the exam. The North Carolina EOG assessments in math, science, and reading all have Cronbach alphas of 0.88 or higher across all forms of each exam[7] (North Carolina Department of Public Instruction, 2014), suggesting the test measures have high internal consistency and reliability.

All models contain a vector of school-level covariates intended to capture school-specific effects that may bias the effect estimate for students who transfer schools or because these variables are correlated with race match due to high rates of racial segregation as described above. Schoolwide covariates include school proficiency rate,[8] percentage of students who are minorities,[9] percentage of students classified as economically disadvantaged (defined as qualifying for free or reduced price meals), a quadratic function of school enrollment measured as average daily membership (average student enrollment over all days of the school year), and a quadratic function of per pupil expenditures (PPE). I also include a set of urbanicity dummies based on the four broad U.S. Census urban classifications: city, suburb, town, and rural.

Some teachers may give higher average grades to all students and others may give lower average grades. If these teacher-specific tendencies are correlated with teacher race or student-teacher race match, they may bias the effect estimate. I control for teacher grading stringency with the mean numerical grade assigned by the teacher, excluding the observed student's course grade. Finally, I control for teacher gender, race, and a vector of experience dummies.

[7] There are multiple parallel test forms for each exam.

[8] School proficiency rate reflects the percent of all EOC and/or EOG exams administered in the school that were passed with a score of proficient or above. Proficiency cut scores are set at the state level.

[9] Students who are Black, Hispanic, Asian, Pacific Islander, American Indian, two or more races, or "other" count toward the school's minority percent.

Empirical strategy

Because students are not randomly assigned to teachers, it is difficult to disentangle the effect of having a race-matched teacher from the mechanisms driving classroom assignment. If principals consider student and teacher race in making classroom assignments, an OLS estimate could be biased. In particular, to the extent that unobserved student characteristics are associated with assignment to a race-matched teacher, the race match estimate will be biased. I attempt to mitigate bias caused by the endogenous race match variable through the use of student, student-by-year, and student-by-subject fixed effects over three years of coursetaking and exams in North Carolina. Specifically, I estimate a series of fixed effects models to predict the effect of student-teacher race match on course grade, conditional on standardized exam scores. Other literature on student-teacher race congruence also uses a student fixed effects approach to address the possible endogeneity of the race congruence variable, but these studies rely only on across-year rather than across-year and across-subject variation within students (e.g., Dee, 2005; Ouazad, 2014). The primary benefit of the longitudinal approach over the within-year approach is that longer panels provide more opportunity for an individual student to have different values of the race match variable across courses, while a possible limitation is more potential to violate the strict exogeneity assumption for fixed effects estimators, which requires errors to be exogenous across time in addition to contemporaneously. Focusing on within-year variation may reduce the possibility of violating this assumption if unobserved student-level characteristics related to teacher assignment and course grade are less likely to vary across courses within a single year than across years within a single class. However, exploiting across-subject variation within a single year would lead to biased estimates if, for example, students were more likely to be race matched in course subjects where they tended to exert more effort. Through the use of both within- and between-year estimators, I am able to compare estimates across different sources of variation. Similar effect estimates across the between-year and between-subject approaches would mitigate concerns about bias. Differences would suggest at least one of the fixed effects models fails to account for one of these unobserved mechanisms, for example, if unobserved student-level characteristics related to student-teacher race match and course grade vary by year or subject.

A final concern about the fixed effects approach is it requires observations within the panel to switch in terms of student-teacher race match. As a result, the race match coefficients represent the effects of race match within the group of switchers rather than the full population. These switchers are defined differently by specification. In the student fixed effects models, switchers are students who have at least one race-matched and one non-race-matched teacher across all courses taken in the three years of data. In the student-by-year fixed effects models, switchers have a race-matched and unmatched teacher within a single year; in other words, these estimates rely on within-year variation. Finally, in the student-by-subject fixed effects models, switchers have a race-matched and unmatched teacher within a single subject area over three years.

Students in the fixed effects estimation samples—i.e., the students for whom there is variation in the race match variable—tend to come from schools that are lower performing, higher minority, and higher poverty than the average school in North Carolina. Estimates should therefore be considered to reflect the experience of students in these schools and may not capture the effect of race match among students in whiter, more affluent schools. Additionally, the student composition of the estimation sample for the student-by-subject fixed effects models is different from the other two estimation sample on two measures; first, sixth- and seventh-graders are overrepresented, with about half of the analytic sample in these grades compared with about one-third of the other two fixed effects sample. Second, reading is overrepresented and science is underrepresented. Therefore, dissimilar estimates in the student-by-subject fixed effects sample could reflect heterogeneous effects of race match by grade or subject. [10]

While the fixed effects framework is intended to control for the bias associated with student-teacher sorting and controlling for exam score partials out the student's subject-matter knowledge as measured by the state exam, two additional factors may affect course grade beyond a student's mastery of the subject. First, teachers may base grades on student activities and behaviors that are not related to exam performance (Farkas et al., 1990). Second, teachers may observe EOG exam scores before assigning grades and use those scores as part of the course grade calculation. In terms of the first factor, student-by-year fixed effects will account for student-specific dynamics that do not vary across subject areas, but will not address student-specific dynamics that vary from course to course or heterogeneous teacher response to unobserved student characteristics or behaviors. Student-by-subject fixed effects will account for unobserved student characteristics that are consistent within subject area, but will not account for time-varying student characteristics or behaviors. In the latter case, if the course-varying (in the case of student-by-year fixed effects) or time-varying (with student-by-subject fixed effects) omitted variable is correlated systematically with both race match and the grade a teacher assigns, the fixed effects estimates would be biased. Similar estimates across model specifications would abate the concern that omitted variable bias is dramatically distorting the results. I also include a rich set of covariates in each of the models in an effort to account for as many explanatory variables as possible. The second factor, that teachers may observe EOC and

[10] A comparison of the estimation samples across the three analytic approaches can be found Appendix Table A-2.

EOG scores prior to assigning course grades, would attenuate the effects of the race match variables because I control for exam score. The student fixed effects model takes the form

$$\begin{aligned} CourseGrade_{ijst} &= \beta_1 RaceMatch_{ijst} + \beta_2 TestScore_{ist} + \alpha \boldsymbol{X'}_{k(t)} + \lambda \boldsymbol{Z'}_{j(t)} + \gamma_t \\ &+ \rho_s + \sigma_i + u_{ijst} \end{aligned} \tag{1}$$

where the course grade for student *i* assigned to teacher *j* for subject *s* in year *t* is a linear function of *RaceMatch*, an indicator equal to 1 if the student and teacher are the same race; *TestScore*, the student's EOG or EOC score; ***X'***, a vector of school characteristics, ***Z'***, a vector of teacher characteristics; γ, a year fixed effect with 2014-15 as the reference category; ρ, a subject fixed effect with math as the reference category; σ, a student fixed effect; and an idiosyncratic error term clustered at the student level. The vector of school characteristics includes school proficiency rate, minority, and economically disadvantaged percentages, per pupil expenditures (PPE) and PPE squared, average daily membership (ADM) and ADM squared, and urbanicity with rural as the reference category. Teacher variables include a series of race dummies with White as the reference category, the mean grade assigned by that teacher for all students other than student *i*, a male teacher dummy, and a vector of teacher experience indicators (1 year of experience, 2 years, 3 years, 4-9 years, more than 10 years) with 0 years as the reference category. The coefficient of interest is β_1, providing the estimated effect of having a race-matched teacher, conditional on EOC or EOG score, year and subject fixed effects, and teacher and school covariates. The student-by-year and student-by-subject fixed effects models follow the same form, except the student-by-year model has no subject fixed effect and the student-by-subject model has no year fixed effect since those slopes vary randomly by student. In each model, I cluster standard errors at the level of the fixed effect.

To answer the second research question, which considers heterogeneous treatment effects based on student race, I generate a series of race-specific race match indicators for each of the race groups that take the value of 1 if a student is both race matched and in that race group and 0 otherwise. The student fixed effects model takes the form

$$\begin{aligned} CourseGrade_{ijst} &= \beta_1 RaceMatch \times White_{ijst} + \beta_2 RaceMatch \times Black_{ijst} \\ &+ \beta_3 RaceMatch \times Hispanic_{ijst} + \beta_4 RaceMatch \times Asian_{ijst} \\ &+ \beta_5 RaceMatch \times AmericanIndian_{ijst} + \beta_6 TestScore_{ist} \\ &+ \alpha \boldsymbol{X}'_{k(t)} + \lambda \boldsymbol{Z}'_{j(t)} + \gamma_t + \rho_s + \sigma_i + u_{ijst} \end{aligned} \tag{2}$$

In these models, β_1 represents the estimated effect of race match for a White student for classes in which she is assigned to a White teacher relative to classes in which she is assigned to a non-White teacher, while β_2 through β_5 follow analogous interpretations for Black, Hispanic, Asian, and American Indian students.

A noteworthy limitation of this analysis is the use of the course grade variable as an outcome. First, because course grades are coded differently across schools, I simplify all numeric grades to A-F grades in order to align them with letter grades across schools. If there is classical measurement error in the simplified outcome variable, the effect estimates will still be unbiased and consistent, although the standard errors will be larger (Wooldridge, 2009). However, if the measurement error is systematically related to one of the covariates, the effect estimates may be biased. Second, the large number of fixed effects require that I treat the ordinal test score outcome as continuous. Given these two methodological limitations, I check the robustness of my effect estimates in two ways. First, I estimate a series of linear probability models to predict the binary outcomes of receiving a course grade of *A*, *B* or higher, *C* or higher, and *D* or higher. If the letter grade estimates are unbiased and the linear probability fixed effects approach is valid, these results should be monotonically similar across each of the four linear probability outcomes, and qualitatively similar to the main results. Second, I reestimate the original models on the subset of observations that have numeric course grades in the administrative dataset using the continuous course grade as the outcome. While this limited sample would not have as strong external validity as a primary analysis, similar results using this approach would suggest the categorical nature of the letter grade in the outcome in the primary analysis did not systematically bias the results.[11] The results from both of these robustness checks are provided in the appendix.

Findings

Table 3, row 1, shows that students receive slightly higher course grades when they are paired with same-race teachers than when they are paired with different-race teachers. Estimates are stable across all three levels of fixed effects, showing that having a race-matched teacher is associated with a course grade increase of .015 to .022 grade points, conditional on school and teacher covariates.[12] These effects represent 1.3 to 2 percent of a standard deviation on the course grade variable. The estimated effects on race match are attenuated slightly with the inclusion of school-level covariates in the models that leverage year-to-year variation across students (columns 2 and 6), underscoring the role of school context in teacher grading practices and the extent to which school context intersects with student-teacher race match. For example, Black students are more likely to be race-matched in lower performing, higher minority schools because these schools have more Black teachers. White students are more likely to be race-matched in higher performing,

[11] I also implement a similar approach in which I extrapolate continuous course grades from the records that include only letter grades as the midpoint of the score range for that grade on a 100-point scale. Results were very similar and are available upon request.

[12] Full results from Table 3 including all estimated coefficients are presented in Appendix Table A-3.

lower minority schools where there are more White teachers (Appendix Table A-1 provides descriptive statistics on school demographics and performance by teacher race).

Table 3 ABOUT HERE

If the effect of student-teacher race match is moderated by student race, the overall estimate may over- or understate the effect for certain race groups. The second research question explores the extent to which these effects vary for students of different races by introducing interactions into the models above. Table 4 provides the estimated effect of race match by student race.[13] The race match effect for Black students, shown in row 2, is consistently positive and significant across all models, with estimates ranging from .036 to .046. The estimates are stronger for Black students than they are in the overall population as measured by the main effects model, representing an effect size of 3.2 to 3.7 percent of a standard deviation.

It is important to note that these estimates cannot speak to the mechanism through which that differential operates; it is possible that Black students receive higher course grades from Black teachers (which might occur in the presence of in-group bias), but it is also possible that Black students receive lower course grades from teachers who are not Black (which might occur in the presence of passive teacher effects such as role model effects or stereotype threat, or active teacher effects such as teacher biases against students who do not share their race or ethnicity), or that the effect is driven by some combination of both active and passive teacher effects. Again, estimates that leverage year-to-year variation are attenuated after controlling for school-level covariates, highlighting the different school settings in which Black students tend to be race matched.

[13] Full results from Table 4 including all estimated coefficients are presented in Appendix Table A-4.

I do not find consistent evidence across specifications that students of other races receive different course grades when they have race-matched teachers. However, the estimates on Hispanic race match, shown in row 3, are positive and significant in the student-by-year models, providing some suggestive evidence that within a given school year, Hispanic students receive higher course grades from Hispanic teachers than from teachers who are not Hispanic. The estimates for Hispanic students are not consistent across specifications, which may stem from the relatively small number of Hispanic students in the sample and the rarity of race match for these students. For example, the Hispanic students with both race matched and non-race matched teachers in a given year may look different than Hispanic students with race matched and non-race matched teachers in a given subject.

The estimates on race match for Asian students, shown in row 4, are consistently positive but insignificant. Of the five race groups in the sample, Asian students are the smallest group and they are race matched the least frequently (approximately 1% of the time, shown in Table 2 above). To that end, the positive estimates suggest the possibility that student-teacher race match may be associated with higher course grades for Asian students, but North Carolina is not the ideal setting to examine that dynamic.

Across non-Black groups, estimates are largely consistent with and without the inclusion of school covariates, shown in odd and even columns of Table 4, respectively. This pattern is different from the pattern of estimates for Black students, for whom school context seems to play a role in the race match effect. The attenuated estimate for Asian students in Column 6 hints at the potential that school context may play a role in the effect of race match for Asian students, but these estimates are too imprecise to draw conclusions.

Table 4 ABOUT HERE

The primary findings are supported by both robustness checks, with results provided in the appendix. On average, students are more likely to receive higher grades in race matched conditions than non-race matched conditions (Appendix Table A-5). Race match is also associated with a higher 0–100 course grade for students in schools that report continuous course grades (Appendix Table A-7). Positive effects of race match for black students were consistent across the two robustness checks. In particular, the probability of receiving an A was slightly less than 1 percentage point higher for Black students when they were race matched than when they were not race matched, and the probability of receiving at least a B or C was more than 1 percentage point higher (Appendix Table A-6). On the continuous course grade outcome, race match is associated with a .24 to .46 percentile point increase on the 100 point course grade (Appendix Table A-8). In standard deviation units (.017 to .033 standard deviations), these estimates are very similar to the letter grade estimates. The robustness checks provide additional evidence of positive effects for Hispanic students as well, with the student fixed effects and student-by-year fixed effects models finding Hispanic students receive a higher continuous course grade when they are race matched and that they have a higher probability of receiving at least a B or at least a C. The models estimating continuous course grade do not find even suggestive positive effects for Asian students, though the small number of Asian students in this sample limit the external validity of the estimates.

Conclusion

This analysis adds to the emerging evidence that students with race-matched teachers are assessed more positively than their peers with different-race teachers, and that the benefits are strongest and most consistent for Black students. While the present study does not speak to the mechanisms behind these effects, the findings are consistent with the active teacher effect hypothesis, where White teachers' biases lead them to assign lower course grades to Black students. In-group bias could also explain these effects if, for example, White teachers simply assigned higher grades to White students rather than systematically assigning lower grades to Black students. The findings could also be consistent with passive teacher effects in which Black students exert more effort for Black teachers (or less effort for non-Black teachers), though only to the extent that the increased or decreased effort did not translate to a higher or lower exam score.

The effect size is small, at about 2 percent of a standard deviation across all students and about 3.5 percent of a standard deviation for Black students. However, the mechanisms driving the course grade outcome may also be steering the test score covariate in the same direction.[14] For example, if teacher perceptions of students affect test scores in the same direction as course grades (as found, e.g., in Cooper, Findley, & Good, 1982; Dee, 2004; Egalite, Kisida, & Winters, 2015; Oates, 2003), then small negative effects of a homogenous teaching workforce will magnify over multiple outcomes. In addition to contributing to the research base about the effects of student-teacher race match on student outcomes for Black students, these findings also present some evidence that having a race-matched teacher can benefit Hispanic students—a topic that is less prevalent in the existing literature, which focuses primarily on Black students.

In sum, this paper contributes to the mounting evidence on the small positive effects of race match on student outcomes, suggesting training and recruiting a more diverse teacher

[14] In fact, estimating the same set of models in this sample with test scores on the left-hand side finds similar results to Egalite et al (2015), with covariate-adjusted main effects on race match at.01 to .016 standard deviations and effects for Black students at about .01 to .016 standard deviations. Like Egalite and colleagues, I also find small positive effects on test scores for race matched white students. Results from these models are presented in Appendix tables A-9 and A-10.

workforce would benefit students—Black students in particular—in multiple ways, including higher course grades and test scores.

References

Adair, A. V. (1984). *Desegregation: The Illusion of Black Progress*. University Press of America.

Bates, L. A., & Glick, J. E. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Social Science Research*, *42*(5), 1180–1190. https://doi.org/10.1016/j.ssresearch.2013.04.005

Bertrand, M., & Mullainathan, S. (2001). Do People Mean What They Say? Implications for Subjective Survey Data. *The American Economic Review*, *91*(2), 67–72.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Chapter 59—Measurement Error in Survey Data. In J. J. H. and E. Leamer (Ed.), *Handbook of Econometrics* (Vol. 5, pp. 3705–3843). Elsevier. http://www.sciencedirect.com/science/article/pii/S1573441201050127

Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology*, *76*(2), 236–247. http://dx.doi.org/10.1037/0022-0663.76.2.236

Cooper, H. M., Findley, M., & Good, T. (1982). Relations between student achievement and various indexes of teacher expectations. *Journal of Educational Psychology*, *74*(4), 577–579. http://dx.doi.org/10.1037/0022-0663.74.4.577

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

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

Downey, D. B., & Pribesh, S. (2004). When Race Matters: Teachers' Evaluations of Students' Classroom Behavior. *Sociology of Education*, *77*(4), 267–282.

Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, *45*, 44–52. https://doi.org/10.1016/j.econedurev.2015.01.007

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial & Labor Relations Review*, *48*(3), 547–561. https://doi.org/10.1177/001979399504800312

Evans, M. O. (1992). An Estimate of Race and Gender Role-Model Effects in Teaching High School. *The Journal of Economic Education*, *23*(3), 209–217. https://doi.org/10.1080/00220485.1992.10844754

Farkas, G., Grobe, R. P., Sheehan, D., & Shuan, Y. (1990). Cultural Resources and School Success: Gender, Ethnicity, and Poverty Groups within an Urban School District. *American Sociological Review*, *55*(1), 127–142. https://doi.org/10.2307/2095708

Feld, J., Salamanca, N., & Hamermesh, D. S. (2016). Endophilia or Exophobia: Beyond Discrimination. *The Economic Journal*, *126*(594), 1503–1527. https://doi.org/10.1111/ecoj.12289

Foster, M. (1993). Educating for Competence in Community and Culture Exploring the Views of Exemplary African-American Teachers. *Urban Education*, *27*(4), 370–394. https://doi.org/10.1177/0042085993027004004

Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, *52*, 209–224. https://doi.org/10.1016/j.econedurev.2016.03.002

Henry, G. T., Barrett, N., & Marder, C. (2016). "Double-Dosing" in Math in North Carolina Public Schools. REL 2016-140. In *Regional Educational Laboratory Southeast*. Regional Educational Laboratory Southeast. https://eric.ed.gov/?id=ED567754

Hess, F. M., & Leal, D. L. (1997). Minority Teachers, Minority Students, and College Matriculation: A New Look at the Role-Modeling Hypothesis. *Policy Studies Journal*, *25*(2), 235–248. https://doi.org/10.1111/j.1541-0072.1997.tb00018.x

Jansson, J., & Tyrefors, B. (2018). *Gender grading bias at Stockholm University: Quasi-experimental evidence from an anonymous grading reform* (IFN Working Paper No. 1226).

Joshi, E., Doan, S., & Springer, M. G. (2018). Student-Teacher Race Congruence: New Evidence and Insight From Tennessee. *AERA Open*, *4*(4), 2332858418817528. https://doi.org/10.1177/2332858418817528

Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, *57*(3), 469–480. http://dx.doi.org/10.1037/0022-3514.57.3.469

Jussim, L., Eccles, J., & Madon, S. (1996). Social Perception, Social Stereotypes, and Teacher Expectations: Accuracy and the Quest for the Powerful Self-Fulfilling Prophecy. *Advances in Experimental Social Psychology*, *28*, 281–388. https://doi.org/10.1016/S0065-2601(08)60240-3

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, *9*(2), 131–155.

Kozlowski, K. P. (2015). Culture or teacher bias? Racial and ethnic variation in student–teacher effort assessment match/mismatch. *Race and Social Problems*, *7*(1), 43–59.

Lavy, V., Sand, E., & Shayo, M. (2018). *Charity Begins at Home (and at School): Effects of Religion-Based Discrimination in Education* (No. w24922; p. w24922). National Bureau of Economic Research. https://doi.org/10.3386/w24922

Loeb, S., & Reininger, M. (2004). Public Policy and Teacher Labor Markets. What We Know and Why It Matters. *Education Policy Center*.

McGrady, P. B., & Reynolds, J. R. (2013). Racial Mismatch in the Classroom: Beyond Black-white Differences. *Sociology of Education*, *86*(1), 3–17.

Oates, G. L. St. C. (2003). Teacher-Student Racial Congruence, Teacher Perceptions, and Test Performance*. *Social Science Quarterly*, *84*(3), 508–525. https://doi.org/10.1111/1540-6237.8403002

Ouazad, A. (2014). Assessed by a teacher like me: Race and teacher assessments. *Education*, *9*(3), 334–372.

Quillian, L. (2008). Does unconscious racism exist? *Social Psychology Quarterly*, *71*(1), 6.

Rocque, M. (2010). Office Discipline and Student Behavior: Does Race Matter? *American Journal of Education*, *116*(4), 557–581. https://doi.org/10.1086/653629

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Takei, Y., & Shouse, R. (2008). Ratings in Black and White: Does racial symmetry or asymmetry influence teacher assessment of a pupil's work habits? *Social Psychology of Education*, *11*(4), 367–387. https://doi.org/10.1007/s11218-008-9064-0

van Ewijk, R. (2011). Same work, lower grade? Student ethnicity and teachers' subjective assessments. *Economics of Education Review*, *30*(5), 1045–1058. https://doi.org/10.1016/j.econedurev.2011.05.008

Vinopal, K. (2020). Socioeconomic Representation: Expanding the Theory of Representative Bureaucracy. *Journal of Public Administration Research and Theory*, *30*(2), 187–201. https://doi.org/10.1093/jopart/muz024

Tables

Table 1. Course grade frequencies and means

	Freq.	Pct.
A	939,377	29.3
B	1,079,614	33.67
C	706,792	22.04
D	331,318	10.33
F	149,345	4.66
Total	*3,206,446*	*100.0*

	Mean	SD
Numerical grade (4-point scale)	2.726	1.127

Numeric grades are on standard 4-point scale, where *A*=4, *B*=3, *C*=2, *D*=1, and *F*=0

Table 2. Descriptive statistics on race match and course grades

	Race match proportion by student race	Mean course grade by race match and student race			*N*
	Proportion race matched	Race matched	Not race matched	All	
White	0.91	2.997	2.811	2.980	1,693,374
	(0.290)	(1.040)	(1.102)	(1.048)	
Black	0.27	2.179	2.356	2.308	841,446
	(0.443)	(1.115)	(1.131)	(1.130)	
Hispanic	0.02	2.426	2.505	2.503	536,145
	(0.123)	(1.152)	(1.139)	(1.139)	
Asian	0.01	3.331	3.277	3.278	89,810
	(0.111)	(0.937)	(0.953)	(0.952)	
American Indian	0.32	2.595	2.519	2.543	45,671
	(0.466)	(1.072)	(1.108)	(1.097)	
Total	0.56	2.888	2.523	2.726	
	(0.497)	(1.086)	(1.145)	(1.127)	
		1,786,125	1,420,321	3,206,446	

NOTE: Student-course observations. Standard deviations in parentheses. Course grades on 4.0 scale.

Table 3. Main effects of student-teacher race match on course grade

	(1) Student FE	(2)	(3) Student-by-year	(4)	(5) Student-by-subject	(6)
Race match	0.023***	0.018***	0.023***	0.022***	0.022***	0.015***
	(0.0016)	(0.0015)	(0.0018)	(0.0018)	(0.0023)	(0.0022)
Test score	0.197***	0.203***	0.142***	0.142***	0.137***	0.147***
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0013)	(0.0013)
Teacher controls	X	X	X	X	X	X
School controls		X		X		X
Year FE	X	X			X	X
Subject FE	X	X	X	X		
Adjusted R2	0.706	0.710	0.755	0.755	0.691	0.698
Adjusted within R2	0.215	0.226	0.165	0.165	0.251	0.269
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies, and the mean teacher grade for that year. Appendix Table A-3 contains full results with all coefficient estimates.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Race match effect by student race

	(1) Student FE	(2)	(3) Student by year	(4)	(5) Student by subject	(6)
Race match white	0.005 (0.0034)	0.001 (0.0033)	0.005 (0.0039)	0.005 (0.0039)	0.003 (0.0049)	-0.003 (0.0048)
Race match Black	0.045*** (0.0036)	0.038*** (0.0035)	0.041*** (0.0041)	0.041*** (0.0041)	0.046*** (0.0052)	0.036*** (0.0051)
Race match Hispanic	0.015 (0.0108)	0.018 (0.0107)	0.033** (0.0122)	0.033** (0.0122)	-0.007 (0.0166)	-0.008 (0.0162)
Race match Asian	0.039 (0.0202)	0.038 (0.0201)	0.017 (0.0242)	0.017 (0.0242)	0.029 (0.0279)	0.012 (0.0281)
Race match American Indian	-0.021 (0.0121)	-0.020 (0.0120)	0.006 (0.0141)	0.005 (0.0141)	-0.017 (0.0180)	-0.011 (0.0179)
Teacher controls	X	X	X	X	X	X
School controls		X		X		X
Year FE	X	X			X	X
Subject FE	X	X	X	X		
Adjusted R^2	0.706	0.710	0.755	0.755	0.691	0.698
Adjusted within R^2	0.215	0.226	0.165	0.165	0.251	0.269
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies, and the mean teacher grade for that year.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix

Table A-1. School-level student demographics by teacher race

	White	Black	Hispanic	Asian	American Indian	Total
Minority percent	44.2 (23.07)	73.1 (21.02)	59.7 (24.71)	58.5 (23.68)	81.5 (21.00)	49.2 (25.24)
White student percent	55.8 (23.06)	26.9 (21.02)	40.3 (24.71)	41.5 (23.67)	18.5 (21.00)	50.8 (25.24)
Black student percent	21.3 (18.07)	46.2 (21.39)	30.7 (20.60)	31.6 (20.69)	22.4 (18.72)	25.2 (20.69)
Hispanic student percent	15.5 (11.17)	18.5 (12.66)	21.5 (15.46)	17.4 (10.98)	12.1 (11.38)	16.0 (11.54)
Asian student percent	2.5 (4.00)	2.7 (3.60)	2.6 (3.59)	4.3 (6.41)	0.7 (1.41)	2.5 (3.95)
American Indian student percent	0.9 (3.77)	1.7 (6.10)	0.6 (1.87)	0.9 (5.40)	42.2 (34.29)	1.4 (6.75)
Performance Composite	57.6 (15.18)	45.9 (16.47)	53.9 (16.75)	55.0 (17.85)	41.4 (11.66)	55.7 (16.02)
Observations	99684					

NOTE: Table cells contain means with standard deviations in parentheses.

Table A-2. Estimation sample descriptives

	Analytic sample	Student FE	Student by year	Student by subject
School				
Minority percent	49.55 (25.262)	61.73 (23.496)	63.58 (22.862)	62.87 (23.199)
Performance Composite	56.26 (15.552)	51.70 (15.901)	50.68 (15.591)	51.17 (15.718)
Economically disadvantaged percent	50.09 (19.002)	54.19 (18.656)	55.39 (18.300)	55.10 (18.458)
Per pupil expenditures	87.27 (21.152)	89.08 (23.863)	89.79 (24.971)	89.08 (23.464)
Enrollment	8.38 (4.667)	8.25 (4.451)	7.90 (4.069)	7.88 (4.015)
City	0.36 (0.479)	0.46 (0.498)	0.46 (0.498)	0.46 (0.499)
Suburb	0.08 (0.267)	0.04 (0.188)	0.03 (0.176)	0.04 (0.184)
Town	0.07 (0.261)	0.08 (0.269)	0.08 (0.273)	0.08 (0.267)
Rural	0.49 (0.500)	0.43 (0.495)	0.43 (0.495)	0.42 (0.494)
Teacher				
Male teacher	0.17 (0.379)	0.18 (0.383)	0.18 (0.386)	0.16 (0.369)
Teacher race: White	0.82 (0.386)	0.61 (0.487)	0.51 (0.500)	0.51 (0.500)
Teacher race: Black	0.15 (0.358)	0.32 (0.465)	0.40 (0.490)	0.41 (0.492)
Teacher race: Hispanic	0.01 (0.110)	0.03 (0.170)	0.04 (0.204)	0.04 (0.189)

	Analytic sample	Student FE	Student by year	Student by subject
Teacher race: Asian	0.01 (0.094)	0.02 (0.134)	0.03 (0.157)	0.02 (0.150)
Teacher race: American Indian	0.01 (0.104)	0.02 (0.143)	0.02 (0.155)	0.02 (0.145)
Mean teacher grade	2.72 (0.610)	2.58 (0.621)	2.57 (0.599)	2.57 (0.620)
Teaching experience				
0 years	0.05 (0.216)	0.06 (0.234)	0.06 (0.237)	0.06 (0.233)
1 year	0.06 (0.237)	0.07 (0.257)	0.08 (0.263)	0.07 (0.257)
2 years	0.06 (0.238)	0.07 (0.248)	0.07 (0.248)	0.07 (0.247)
3 years	0.06 (0.229)	0.06 (0.231)	0.05 (0.227)	0.06 (0.230)
4-9 years	0.25 (0.431)	0.25 (0.431)	0.25 (0.431)	0.25 (0.432)
10+ years	0.53 (0.499)	0.50 (0.500)	0.50 (0.500)	0.50 (0.500)
Student race				
White	0.52 (0.499)	0.47 (0.499)	0.46 (0.498)	0.46 (0.498)
Black	0.27 (0.442)	0.47 (0.499)	0.48 (0.500)	0.49 (0.500)
Hispanic	0.17 (0.372)	0.03 (0.183)	0.03 (0.179)	0.03 (0.160)
Asian	0.03 (0.166)	0.00 (0.068)	0.00 (0.065)	0.00 (0.063)

	Analytic sample	Student FE	Student by year	Student by subject
American Indian	0.01 (0.118)	0.02 (0.134)	0.02 (0.140)	0.02 (0.132)
Student grade				
Grade 5	0.20 (0.401)	0.15 (0.355)	0.13 (0.341)	0.12 (0.322)
Grade 6	0.16 (0.364)	0.17 (0.372)	0.15 (0.357)	0.24 (0.425)
Grade 7	0.16 (0.363)	0.19 (0.390)	0.16 (0.366)	0.26 (0.440)
Grade 8	0.25 (0.435)	0.31 (0.464)	0.42 (0.494)	0.27 (0.444)
Grade 9	0.07 (0.248)	0.06 (0.243)	0.02 (0.128)	0.05 (0.211)
Grade 10	0.15 (0.353)	0.11 (0.315)	0.12 (0.319)	0.06 (0.246)
Subject area				
Math	0.34 (0.473)	0.36 (0.480)	0.35 (0.477)	0.40 (0.491)
Reading	0.41 (0.492)	0.41 (0.493)	0.40 (0.491)	0.52 (0.499)
Science	0.25 (0.432)	0.23 (0.419)	0.25 (0.432)	0.07 (0.260)
Observations	3,515,909	1,083,810	536,792	425,603

Table A-3. Full results from Table 3 estimating race match main effect on course grade

	(1)	(2)	(3)	(4)	(5)	(6)
	Student FE		Student-by-year		Student-by-subject	
Race match	0.023***	0.018***	0.023***	0.022***	0.022***	0.015***
	(0.0016)	(0.0015)	(0.0018)	(0.0018)	(0.0023)	(0.0022)
Test score	0.197***	0.203***	0.142***	0.142***	0.137***	0.147***
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0013)	(0.0013)
Reading	0.008***	0.003**	0.026***	0.026***		
	(0.0009)	(0.0009)	(0.0009)	(0.0009)		
Science	0.006***	0.004***	0.036***	0.036***		
	(0.0010)	(0.0010)	(0.0010)	(0.0010)		
2016	-0.076***	-0.063***			-0.040***	-0.013***
	(0.0012)	(0.0013)			(0.0012)	(0.0013)
2017	-0.097***	-0.079***			-0.049***	-0.007***
	(0.0014)	(0.0018)			(0.0014)	(0.0018)
Male teacher	0.030***	0.027***	0.024***	0.024***	0.031***	0.026***
	(0.0012)	(0.0012)	(0.0014)	(0.0014)	(0.0018)	(0.0018)
Mean teacher grade	0.668***	0.694***	0.672***	0.673***	0.719***	0.756***
	(0.0013)	(0.0013)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
1 year experience	-0.028***	-0.025***	-0.019***	-0.019***	-0.025***	-0.018***
	(0.0027)	(0.0027)	(0.0032)	(0.0032)	(0.0040)	(0.0039)
2 years experience	-0.035***	-0.030***	-0.018***	-0.018***	-0.035***	-0.026***
	(0.0027)	(0.0027)	(0.0031)	(0.0031)	(0.0040)	(0.0039)
3 years experience	-0.040***	-0.034***	-0.025***	-0.025***	-0.038***	-0.029***
	(0.0028)	(0.0027)	(0.0032)	(0.0032)	(0.0041)	(0.0040)
4-9 years experience	-0.050***	-0.041***	-0.031***	-0.031***	-0.049***	-0.035***
	(0.0022)	(0.0022)	(0.0026)	(0.0026)	(0.0033)	(0.0032)
10+ years experience	-0.063***	-0.053***	-0.036***	-0.036***	-0.065***	-0.049***
	(0.0022)	(0.0021)	(0.0025)	(0.0025)	(0.0032)	(0.0031)
Black	0.025***	0.009***	0.001	0.001	0.040***	0.014***
	(0.0015)	(0.0015)	(0.0017)	(0.0017)	(0.0021)	(0.0021)
Hispanic	0.007	-0.001	-0.003	-0.003	0.023***	0.010
	(0.0043)	(0.0043)	(0.0049)	(0.0049)	(0.0066)	(0.0065)
Asian	0.010*	0.007	0.013*	0.013*	0.018*	0.015*
	(0.0049)	(0.0049)	(0.0056)	(0.0056)	(0.0074)	(0.0073)

	(1)	(2)	(3)	(4)	(5)	(6)
	Student FE		Student-by-year		Student-by-subject	
American Indian	0.016**	0.005	-0.003	-0.003	0.046***	0.027***
	(0.0051)	(0.0050)	(0.0057)	(0.0057)	(0.0078)	(0.0077)
Minority percent		0.002***		0.001		0.002***
		(0.0001)		(0.0017)		(0.0001)
Performance Composite		-0.010***		-0.019***		-0.011***
		(0.0001)		(0.0023)		(0.0001)
Economically disadvantaged percent		0.001***		-0.001		0.002***
		(0.0001)		(0.0021)		(0.0001)
Per pupil expenditures		0.002***		0.005***		0.002***
		(0.0001)		(0.0014)		(0.0001)
Per pupil expenditures squared		-0.000**		-0.000**		-0.000
		(0.0000)		(0.0000)		(0.0000)
Enrollment		-0.001		-0.004		0.002
		(0.0008)		(0.0144)		(0.0008)
Enrollment squared		0.000***		0.000		-0.000*
		(0.0000)		(0.0005)		(0.0000)
City		-0.078***		-0.039		-0.048***
		(0.0092)		(0.0740)		(0.0092)
Suburb		0.007		0.017		0.006
		(0.0152)		(0.1274)		(0.0149)
Town		0.027*		0.162		0.033*
		(0.0135)		(0.1156)		(0.0131)
Constant	1.000***	1.212***	0.892***	1.611***	0.833***	1.016***
	(0.0044)	(0.0172)	(0.0050)	(0.2619)	(0.0057)	(0.0175)
Adjusted R^2	0.706	0.710	0.755	0.755	0.691	0.698
Adjusted within R^2	0.215	0.226	0.165	0.165	0.251	0.269
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies, and the mean teacher grade for that year.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-4. Full results from Table 4 estimating race match moderated effects on course grade

	(1)	(2)	(3)	(4)	(5)	(6)
	Student FE		Student-by-year		Student by subject	
Race match white	0.005	0.001	0.005	0.005	0.003	-0.003
	(0.0034)	(0.0033)	(0.0039)	(0.0039)	(0.0049)	(0.0048)
Race match black	0.045***	0.038***	0.041***	0.041***	0.046***	0.036***
	(0.0036)	(0.0035)	(0.0041)	(0.0041)	(0.0052)	(0.0051)
Race match Hispanic	0.015	0.018	0.033**	0.033**	-0.007	-0.008
	(0.0108)	(0.0107)	(0.0122)	(0.0122)	(0.0166)	(0.0162)
Race match Asian	0.039	0.038	0.017	0.017	0.029	0.012
	(0.0202)	(0.0201)	(0.0242)	(0.0242)	(0.0279)	(0.0281)
Race match American Indian	-0.021	-0.020	0.006	0.005	-0.017	-0.011
	(0.0121)	(0.0120)	(0.0141)	(0.0141)	(0.0180)	(0.0179)
Test score	0.197***	0.203***	0.142***	0.142***	0.137***	0.147***
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0013)	(0.0013)
Reading	0.008***	0.003**	0.026***	0.026***		
	(0.0009)	(0.0009)	(0.0009)	(0.0009)		
Science	0.006***	0.004***	0.036***	0.036***		
	(0.0010)	(0.0010)	(0.0010)	(0.0010)		
2016	-0.076***	-0.063***			-0.040***	-0.013***
	(0.0012)	(0.0013)			(0.0012)	(0.0013)
2017	-0.097***	-0.079***			-0.049***	-0.007***
	(0.0014)	(0.0018)			(0.0014)	(0.0018)
Male teacher	0.029***	0.027***	0.024***	0.024***	0.031***	0.026***
	(0.0012)	(0.0012)	(0.0014)	(0.0014)	(0.0018)	(0.0018)
Mean teacher grade	0.668***	0.694***	0.672***	0.673***	0.719***	0.756***
	(0.0013)	(0.0013)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
1 year experience	-0.028***	-0.025***	-0.019***	-0.019***	-0.025***	-0.018***
	(0.0027)	(0.0027)	(0.0032)	(0.0032)	(0.0040)	(0.0039)
2 years experience	-0.035***	-0.030***	-0.019***	-0.019***	-0.035***	-0.026***
	(0.0027)	(0.0027)	(0.0031)	(0.0031)	(0.0040)	(0.0039)
3 years experience	-0.040***	-0.034***	-0.025***	-0.025***	-0.038***	-0.029***
	(0.0028)	(0.0027)	(0.0032)	(0.0032)	(0.0041)	(0.0040)

	(1)	(2)	(3)	(4)	(5)	(6)
	Student FE		Student-by-year		Student by subject	
4-9 years experience	-0.050^{***}	-0.041^{***}	-0.031^{***}	-0.031^{***}	-0.049^{***}	-0.035^{***}
	(0.0022)	(0.0022)	(0.0026)	(0.0026)	(0.0033)	(0.0032)
10+ years experience	-0.063^{***}	-0.053^{***}	-0.036^{***}	-0.036^{***}	-0.065^{***}	-0.049^{***}
	(0.0022)	(0.0021)	(0.0025)	(0.0025)	(0.0032)	(0.0031)
Black	0.010^{***}	-0.005	-0.013^{***}	-0.013^{***}	0.023^{***}	-0.002
	(0.0028)	(0.0028)	(0.0032)	(0.0032)	(0.0040)	(0.0040)
Hispanic	0.001	-0.008	-0.013^{*}	-0.013^{*}	0.022^{**}	0.008
	(0.0052)	(0.0051)	(0.0059)	(0.0059)	(0.0079)	(0.0077)
Asian	0.001	-0.002	0.006	0.006	0.009	0.007
	(0.0053)	(0.0052)	(0.0061)	(0.0061)	(0.0080)	(0.0079)
American Indian	0.019^{**}	0.008	-0.006	-0.005	0.048^{***}	0.026^{**}
	(0.0059)	(0.0059)	(0.0066)	(0.0066)	(0.0092)	(0.0090)
Minority percent		0.002^{***}		0.001		0.001^{***}
		(0.0001)		(0.0017)		(0.0001)
Performance Composite		-0.010^{***}		-0.019^{***}		-0.011^{***}
		(0.0001)		(0.0023)		(0.0001)
Economically disadvantaged percent		0.001^{***}		-0.001		0.002^{***}
		(0.0001)		(0.0021)		(0.0001)
Per pupil expenditures		0.002^{***}		0.005^{***}		0.001^{***}
		(0.0001)		(0.0014)		(0.0001)
Per pupil expenditures squared		-0.000^{**}		-0.000^{**}		-0.000
		(0.0000)		(0.0000)		(0.0000)
Enrollment		-0.001		-0.004		0.002
		(0.0008)		(0.0144)		(0.0008)
Enrollment squared		0.000^{***}		0.000		-0.000^{**}
		(0.0000)		(0.0005)		(0.0000)
City		-0.078^{***}		-0.039		-0.048^{***}
		(0.0092)		(0.0740)		(0.0092)
Suburb		0.007		0.017		0.006
		(0.0152)		(0.1274)		(0.0149)
Town		0.027^{*}		0.162		0.033^{*}
		(0.0135)		(0.1156)		(0.0131)

	(1)	(2)	(3)	(4)	(5)	(6)
	Student FE		Student-by-year		Student by subject	
Constant	1.010***	1.222***	0.901***	1.622***	0.844***	1.026***
	(0.0047)	(0.0172)	(0.0053)	(0.2620)	(0.0061)	(0.0177)
Adjusted R^2	0.706	0.710	0.755	0.755	0.691	0.698
Adjusted within R^2	0.215	0.226	0.165	0.165	0.251	0.269
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies, and the mean teacher grade for that year.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-5.1. Linear probability results, student FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A		B or above		C or above		D or above	
Race match	0.008***	0.007***	0.007***	0.005***	0.006***	0.004***	0.002***	0.002***
	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0007)	(0.0007)	(0.0005)	(0.0005)
Test score	0.054***	0.056***	0.071***	0.073***	0.049***	0.051***	0.021***	0.022***
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Teacher controls	X	X	X	X	X	X	X	X
School controls		X		X		X		X
Year FE	X	X	X	X	X	X	X	X
Subject FE	X	X	X	X	X	X	X	X
Adjusted R^2	0.556	0.558	0.556	0.558	0.453	0.456	0.350	0.352
Adjusted within R^2	0.060	0.064	0.111	0.117	0.102	0.107	0.056	0.060
N	3112730	3112730	3112730	3112730	3112730	3112730	3112730	3112730

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, teacher race with white as the reference category, and the mean teacher grade for that year. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-5.2. Linear probability results, student-by-year FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A		B or above		C or above		D or above	
Race match	0.007***	0.007***	0.007***	0.007***	0.005***	0.005***	0.003***	0.003***
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0008)	(0.0008)	(0.0005)	(0.0005)
Test score	0.041***	0.041***	0.053***	0.053***	0.035***	0.035***	0.013***	0.013***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Teacher controls	X	X	X	X	X	X	X	X
School controls		X		X		X		X
Year FE								
Subject FE	X	X	X	X	X	X	X	X
Adjusted R^2	0.595	0.595	0.602	0.602	0.529	0.529	0.476	0.476
Adjusted within R^2	0.052	0.052	0.083	0.084	0.063	0.064	0.025	0.025
N	2668150	2668150	2668150	2668150	2668150	2668150	2668150	2668150

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, teacher race with white as the reference category, and the mean teacher grade for that year. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-5.3. Linear probability results, student-by-subject FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A		B or above		C or above		D or above	
Race match	0.006***	0.005***	0.006***	0.004**	0.008***	0.005***	0.002**	0.001
	(0.0010)	(0.0010)	(0.0012)	(0.0012)	(0.0010)	(0.0010)	(0.0007)	(0.0007)
Test score	0.027***	0.030***	0.045***	0.048***	0.042***	0.045***	0.022***	0.024***
	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(0.0006)	(0.0006)	(0.0004)	(0.0004)
Teacher controls	X	X	X	X	X	X	X	X
School controls		X		X		X		X
Year FE	X	X	X	X	X	X	X	X
Subject FE								
Adjusted R^2	0.548	0.551	0.541	0.545	0.420	0.425	0.249	0.254
Adjusted within R^2	0.067	0.073	0.130	0.138	0.126	0.135	0.077	0.083
N	1834765	1834765	1834765	1834765	1834765	1834765	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, teacher race with white as the reference category, and the mean teacher grade for that year. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table A-6. Linear probability results by race

	Student FE				Student by year				Student by subject			
	A	B or above	C or above	D or above	A	B or above	C or above	D or above	A	B or above	C or above	D or above
Race match white	0.006***	-0.003	-0.001	-0.001	0.007***	-0.002	-0.000	0.000	0.002	-0.004	-0.000	-0.002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)
Race match Black	0.008***	0.014***	0.011***	0.005***	0.008***	0.015***	0.011***	0.007***	0.008***	0.012***	0.011***	0.004*
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
Race match Hispanic	-0.006	0.007	0.010*	0.007*	-0.002	0.016*	0.014*	0.004	-0.005	-0.009	0.004	0.002
	(0.005)	(0.006)	(0.005)	(0.003)	(0.006)	(0.007)	(0.006)	(0.004)	(0.007)	(0.009)	(0.008)	(0.005)
Race match Asian	0.028*	0.018	-0.002	-0.005	0.019	0.017	-0.012	-0.007	-0.006	0.021	0.000	-0.002
	(0.011)	(0.010)	(0.008)	(0.004)	(0.014)	(0.012)	(0.010)	(0.005)	(0.016)	(0.014)	(0.010)	(0.006)
Race match American Indian	-0.001	0.002	-0.014*	-0.007*	0.005	0.018*	-0.014*	-0.004	-0.006	0.006	-0.000	-0.011*
	(0.005)	(0.007)	(0.006)	(0.003)	(0.007)	(0.008)	(0.007)	(0.004)	(0.008)	(0.010)	(0.009)	(0.005)
Test score	0.056***	0.073***	0.051***	0.022***	0.041***	0.053***	0.035***	0.013***	0.030***	0.048***	0.045***	0.024***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Teacher controls	X	X	X	X	X	X	X	X	X	X	X	X
School controls	X	X	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X					X	X	X	X
Subject FE	X	X	X	X	X	X	X	X				
Adjusted R^2	0.558	0.558	0.456	0.352	0.595	0.602	0.529	0.476	0.551	0.545	0.425	0.254
Adjusted within R^2	0.064	0.117	0.107	0.060	0.052	0.084	0.064	0.025	0.073	0.138	0.135	0.083
N	3112730	3112730	3112730	3112730	2668150	2668150	2668150	2668150	1834765	1834765	1834765	1834765

NOTE: Estimates from models with teacher and school covariates.

Table A-7. Race match main effects (outcome=continuous course grade)

	(1) Student FE	(2)	(3) Student-by-year	(4)	(5) Student-by-subject	(6)
Race match	0.023***	0.018***	0.023***	0.022***	0.022***	0.015***
	(0.0016)	(0.0015)	(0.0018)	(0.0018)	(0.0023)	(0.0022)
Test score	0.197***	0.203***	0.142***	0.142***	0.137***	0.147***
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0013)	(0.0013)
Teacher controls	X	X	X	X	X	X
School controls		X		X		X
Year FE	X	X			X	X
Subject FE	X	X	X	X		
Adjusted R^2	0.706	0.710	0.755	0.755	0.691	0.698
Adjusted within R^2	0.215	0.226	0.165	0.165	0.251	0.269
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies, and the mean teacher grade for that year.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-8. Race match effects by student race (outcome=continuous course grade)

	(1) Student FE	(2)	(3) Student by year	(4)	(5) Student by subject	(6)
Race match white	-0.014	-0.041	0.069	0.069	-0.067	-0.113
	(0.0450)	(0.0444)	(0.0393)	(0.0392)	(0.0680)	(0.0665)
Race match Black	0.400***	0.328***	0.463***	0.456***	0.366***	0.238**
	(0.0501)	(0.0493)	(0.0422)	(0.0422)	(0.0751)	(0.0734)
Race match Hispanic	0.328*	0.316*	0.455***	0.460***	-0.144	-0.225
	(0.1440)	(0.1431)	(0.1267)	(0.1266)	(0.2324)	(0.2269)
Race match Asian	-0.362	-0.465	-0.269	-0.267	-0.145	-0.407
	(0.2809)	(0.2819)	(0.2483)	(0.2484)	(0.5275)	(0.5329)
Race match American Indian	-0.979***	-0.962***	-0.278	-0.282	-1.060***	-0.998***
	(0.1381)	(0.1368)	(0.1523)	(0.1521)	(0.2487)	(0.2468)
Teacher controls	X	X	X	X	X	X
School controls		X		X		X
Year FE	X	X			X	X
Subject FE	X	X	X	X		
Adjusted R^2	0.768	0.772	0.853	0.853	0.697	0.707
Adjusted within R^2	0.301	0.314	0.218	0.219	0.365	0.386
N	2608443	2608443	2258024	2258024	1399938	1399938

Table A-9. Race match main effects on standardized test scores

	(1) Student FE	(2)	(3) Student-by-year	(4)	(5) Student-by-subject	(6)
Race match	0.015*** (0.0012)	0.016*** (0.0012)	0.016*** (0.0015)	0.016*** (0.0015)	0.010*** (0.0017)	0.011*** (0.0017)
Teacher controls	X	X	X	X	X	X
School controls		X		X		X
Year FE	X	X			X	X
Subject FE	X	X	X	X		
Adjusted R^2	0.759	0.760	0.748	0.748	0.752	0.753
Adjusted within R^2	0.002	0.006	0.001	0.001	0.001	0.007
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies, and the mean teacher grade for that year.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-10. Race match moderated effects on standardized test scores

	(1) Student FE	(2)	(3) Student by year	(4)	(5) Student by subject	(6)
Race match white	0.015***	0.016***	0.019***	0.019***	0.012***	0.015***
	(0.0025)	(0.0025)	(0.0033)	(0.0033)	(0.0037)	(0.0037)
Race match Black	0.016***	0.016***	0.014***	0.014***	0.008*	0.008*
	(0.0026)	(0.0026)	(0.0034)	(0.0034)	(0.0038)	(0.0037)
Race match Hispanic	-0.004	-0.003	0.007	0.007	-0.001	0.002
	(0.0077)	(0.0077)	(0.0099)	(0.0099)	(0.0118)	(0.0118)
Race match Asian	-0.026	-0.033	0.002	0.002	-0.104***	-0.108***
	(0.0172)	(0.0171)	(0.0219)	(0.0219)	(0.0259)	(0.0258)
Race match American Indian	0.038***	0.038***	0.030*	0.030*	0.026	0.026
	(0.0098)	(0.0097)	(0.0123)	(0.0123)	(0.0145)	(0.0145)
Teacher controls	X	X	X	X	X	X
School controls		X		X		X
Year FE	X	X			X	X
Subject FE	X	X	X	X		
Adjusted R^2	0.759	0.760	0.748	0.748	0.752	0.753
Adjusted within R^2	0.002	0.006	0.001	0.001	0.001	0.007
N	3112730	3112730	2668150	2668150	1834765	1834765

NOTE: Standard errors clustered at the level of the fixed effect. School controls include minority percent, economically disadvantaged percent, proficiency rate, per pupil expenditures and PPE squared, average daily membership and ADM squared, and urbanicity with rural as the reference category. Teacher controls include gender with female as the reference category, a vector of teacher experience variables with 0 as the reference category, a vector of teacher race dummies with white as the reference category, and the mean teacher grade for that year.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$