The Student-Teacher Race Match Effect on Learning Skills and Behavioral Outcomes*

W. Jesse Wood†
Michigan State University
Department of Economics

Updated: November 2021

Abstract

I provide evidence that diversifying the labor supply of teachers to better reflect the racial distribution of students increases learning and behavioral outcomes for students of color without diminishing outcomes for white students. I use administrative data spanning from 2007 to 2017 within the Los Angeles Unified School District, one of the most racially diverse school districts in the country, to measure the effect of student-teacher race matching on various noncognitive outcomes. I mitigate the concern that race matches are endogenous by including school-grade and student fixed effects into a linear regression model. This setting accounts for any potential sorting that occurs across schools with regards to the racial distribution of teachers as well as any unobserved time invariant student characteristics that may be correlated with race matching. Following a similar method from Jackson (2018), I generate a behavior (using suspensions, absences, and grade retention) and a learning skills (using GPA, marks for work habits, and marks for cooperation) index for each student and find that race matched students in grades 6 through 12 are expected to increase in their behavioral index by 0.041 standard deviations and increases in their learning skills index by 0.011 standard deviations. My findings indicate that students of color also experience increases in the individual components of GPA, work habits, and cooperation and see decreases in absenteeism when matched with a teacher of the same race. I do not find statistically significant effects on any of these outcomes for White students. Because noncognitive outcomes lead to higher high school graduation rates, college enrollment rates, and wages, such effects could lead to a tightening in the achievement and wage gap found between students of color and white students. This result can be achieved with an increase in institutional efforts to ensure teacher populations more closely reflect that of their students.

*I am grateful to the Los Angeles Unified School District for providing the data necessary to conduct this research, and in particular to Vivian Ekchian, Sergio Franco, Cynthia Lim, and Emily Mohr for their partnership in developing the research agenda; and Jolene Chavira, Inocencia Cordova, Kathy Hayes, Crystal Jewett, Joshua Klarin, Jonathan Lesser, and Kevon Tucker-Seeley for their assistance in obtaining and understanding the data. I also wish to thank Scott Imberman, Katharine Strunk, Todd Elder, Jeffrey Wooldridge, and Ijun Lai for their continued guidance and support. I thank conference participants at the Association for Public Policy Analysis and Management (APPAM), Associate for Education Finance and Policy (AEFP), Missouri Valley Economic Association (MVEA), Southern Economic Association (SEA), Michigan State Interdisciplinary Training in Education and Social Sciences (MITTENSS) for helpful comments and discussions. Finally, I'd like to thank my loving wife, Anna. This research was supported in part by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations. For more information about CALDER funders, see www. caldercenter.org/about-calder. All opinions expressed in this article are those of the author and do not necessarily reflect the views of our funders or the institutions with which the authors are affiliated.

†E-mail: woodwil2@msu.edu; Website: www.wjessewood.com

1 Introduction

A troubling characteristic of the U.S. education system is the sustained gap in educational outcomes between White and non-White students. While most of the focus has been on test score gaps (Todd and Wolpin, 2007; Reardon and Galindo, 2009; Fryer and Levitt, 2013), recent studies show that gaps in skills not directly measured by test scores also exist between White students and their non-White peers, even when controlling for environmental factors (Elder and Zhou, 2021; Hashim et al., 2018). Moreover, gaps in these skills, known as noncognitive skills, are equally as concerning as achievement (i.e., test score) gaps because they play a significant role in long-run outcomes such as high school graduation, college enrollment, college major choice, employment, and wages (e.g., Price, 2010; Heckman et al., 2012, 2013; Jackson, 2018).

Research has established that test scores, particularly for students of color, improve when students match the race of their teachers (e.g., Dee, 2004; Egalite et al., 2015; Joshi et al., 2018). However, a teacher's ability to improve test scores is weakly related to that teacher's ability to improve noncognitive skills (Petek and Pope, 2021; Liu and Loeb, 2021; Jackson, 2018; Jennings and DiPrete, 2010). For this reason, even though the link between test scores and race matching is empirically established, the link between noncognitive skills and race matching warrant dedicated study. To this end, recent studies suggest that race matching also significantly affects other outcomes. For example, Lindsay and Hart (2017) find decreases in disciplinary actions for race matched Black students by 3 to 9 percent. In addition, Wright et al. (2017) find that teachers reported race matched students exhibiting fewer externalizing behaviors, such as arguing and getting angry, than non-race matched students. Motivated by the racial gap in gifted education, Grissom et al. (2017) find that race matched Black and Latino students are more likely to be recommended for enrollment in these programs. Recently, Harbatkin (2021) found that race-matched Black students receive higher grade point averages, a measure influenced by student behavior and effort (Guskey, 2018), than non-race matched Black students.

This paper contributes to our growing understanding of race match and student noncognitive skills in three critical ways. First, I examine the relationship between student-teacher race matching and multiple non-test score outcomes. Using data from the Los Angeles Unified School District (LAUSD), I study the relationship between a student's race match status on suspensions, absences, grade retention, grade point average (GPA), effort, and cooperation. I find increases in GPA, marks for effort, and marks for cooperation for race-matched Asian, Black, and Latino students. I also find that race matching is associated with decreased grade retention and absenteeism for Asian and Black students.

However, a student's cognitive abilities may influence these outcomes. For example, it is reasonable to expect a student's GPA, usually calculated by their in-class assignments and test scores, to be affected by cognitive skills. To parse out as much of the noncognitive aspects of these measures as possible, I employ a factor analysis to generate two separate indices. Specifically, I distinguish between "learning skills" and "behavioral" noncognitive measures and create an index for each category. The learning skills index measures a student's ability to manage their time, exercise self-control, cooperate with classmates, and resolve conflicts appropriately. The behavior index captures aspects of externalizing behaviors such as aggression and disruptiveness. I find that race-matched secondary school (i.e., grades 6 through 12) students of color experience increases in overall learning skills and behavior.

The second contribution of this paper is that I broaden the scope of race match and noncognitive studies beyond traditionally studied races and settings. Previous studies have focused on the race match effects for Black and White students at the state or national level,[1] even though Latino students make up a quarter of the student population (NCES, 2017). By leveraging the racial composition and size of LAUSD, I find large and significant increases in behavior and learning skills for Asian, Black, and Latino students. These findings shed light on previously understudied groups of students, which is essential in understanding how to close the noncognitive skills gap between them and their White peers. Also, LAUSD is located within an urban setting, which typically have higher proportions of students of color than non-urban settings (Pew Research Center, 2018). If the purpose of race matching is to shrink the achievement and noncognitive gaps between students of color and their White counterparts, it is vital to

[1]One exception is that Shirrell et al. (2021) find race matched Latino students experience a lower likelihood of suspension than their non-race matched counterparts in New York City.

understand the effect that race matching has where there are the most students of color.

Finally, I shed light on the mechanisms through which race matching impacts a student's noncognitive skills. The existing literature hypothesizes that teachers and students may induce race match effects through active or passive channels. I parse out the race match effects attributed to teacher bias by modifying a method proposed by Elder and Zhou (2021). Student-side mechanisms (e.g., stereotype threat or role model effect) are more prominent than teacher-side mechanisms for behavioral outcomes. In contrast, teacher-side race matching mechanisms (e.g., developing culturally relevant pedagogy) have more influence on improving students' learning skills. The following section further describes how race matching mechanisms may impact a student's behavior and learning skills.

2 Theoretical mechanisms for the race match effect

Two mechanisms have emerged from the literature when describing the role of student-teacher race matching on student outcomes: passive channels and active channels. Passive channels rely on a subconscious response to race, while active channels are a more deliberate reaction to race. These channels may work through either student-side mechanisms (i.e., how a student responds to being race matched to their teacher) or teacher-side mechanisms (i.e., how a teacher responds to a student that shares the same race). While the section below describes each channel separately, there is certainly potential for overlap between them.

2.1 Passive channels

Students of color assigned to teachers of a different race may be influenced by unconscious biases in the classroom environment, potentially limiting students' educational opportunities. For example, students of color are more likely to be perceived as disruptive and deserving of harsher punishment when assigned to White teachers (e.g., Okonofua et al., 2016; Skiba et al., 2011). These implicit biases may be the driving factor in the disproportionate rate at which Black and Latino students are disciplined compared to their Asian and White counterparts (Rodriguez, 2020). Evidence indicates that Black teachers have lower levels of implicit biases towards Black students (e.g., Chin et al., 2020), which may explain why race matched Black students have a lower likelihood of being suspended than when placed in a classroom with a White teacher (e.g., Lindsay and Hart, 2017; Shirrell et al., 2021). Furthermore, when students are not race matched, teachers tend to have significantly lower academic expectations for students of color and perceive their abilities as lower than both their white peers as well as their actual abilities (Fox, 2015; Gershenson et al., 2016; Ouazad, 2014; Ready and Wright, 2011; Tenenbaum and Ruck, 2007). Correspondingly, Black students are perceived as having higher academic ability and exhibiting more pro-academic behaviors in years when they are matched with Black teachers, compared to years when they are demographically mismatched with their teachers (Downey and Pribesh, 2004; Ehrenberg et al., 1995; McGrady and Reynolds, 2013; Oates, 2003).

On the student side, being assigned to a teacher of a different race could activate stereotype threat. This theory, introduced by psychologists Claude Steele and Joshua Aronson, suggests that people are often afraid of behaving in ways that confirm negative stereotypes. Furthermore, when these stereotypes are made salient to students, this fear can cause students to significantly underperform in academic settings (i.e., Steele and Aronson, 1997; Steele, 1995). In other words, students may unconsciously react to teachers' (unconscious or conscious) biases by performing poorly in classes. When students are placed with a teacher of the same race, they may be less likely to worry about judging their actions against negative stereotypes. Instead, students of color may feel more comfortable asking questions and seeking help in race matched classrooms. In testing for stereotype threat, Dee (2015) finds an increase in achievement for minority students that undergo a daily 15-minute affirmation exercise. He explains that this exercise reduces the amount of anxiety experienced by students that feel they are subjected to different treatment based on their socioeconomic status. He theorizes that students who feel the pressures of stereotype threat underachieve, particularly on exams. For the case where students of color feel more comfortable with minority teachers, there may be a reduction or elimination of this perceived threat, which could increase educational outcomes for these students, such as exhibiting more cooperation and effort.

2.2 Active channels

Proponents of culturally relevant pedagogy argue that race matched teachers can more easily build upon a shared understanding to develop better relationships with their students. Additionally, teachers of color may be more proactive in developing culturally relevant curriculum and classroom strategies that can improve students' connections to the course materials (Gay, 2000; Milner, 2011; Saft and Pianta, 2001; Ladson-Billings, 1995). Some teachers may also be more likely to exhibit forms of positive bias towards race matched students. For example, race matched students may receive more attention and positive feedback, or teachers may have a more empathetic mindset (i.e., be more willing to give students the benefit of the doubt) when they misbehave in class (Grissom et al., 2015; Lim, 2006; Okonofua et al., 2021).

Race matched teachers could also influence how parents interact with schools. A large amount of literature documents the importance of parental involvement in children's academic and socio-emotional development (see Van Voorhis et al., 2013, for a thorough review). Studies have also recognized lower parental involvement for students of color for myriad reasons (e.g., Lareau, 2011). However, these parents may feel more comfortable reaching out to teachers of color with their questions or concerns. This increased involvement could result in parents advocating for their children's access to more school resources (Harackiewicz et al., 2012; Lareau and Horvat, 1999).

As previously mentioned, studies have also shown that race matched teachers are more likely to have higher academic expectations for students of color (e.g., Gay, 2000). These higher expectations can cause significant shifts in students' behaviors and effort levels (Ferguson, 2003; Grissom et al., 2015). In addition, students of color have reported feeling more motivated and engaged in classroom activities when they are race matched, compared to when assigned to white teachers (Cherng and Halpin, 2016).

Gershenson et al. (2018) describe how a student-teacher *role model effect* may also be the driving mechanism in their finding of an increase in high school graduation for race matched Black students. Also, Gershenson et al. (2016) find that non-Black teachers exhibit lower expectations than Black teachers for the same Black student. If students realize these perceptions, they may actively seek role models to inspire them.

While I am unable to directly separate the effects of active and passive channels, I attempt to parse out teacher-side mechanisms from the race match effect in Section 3.5. Intuitively, the proposed method controls for a teacher's bias. The resulting analysis reveals the effect that student-side mechanisms have on race matching. By directly comparing the results of specifications with teacher bias adjusted outcomes to those without controlling for teacher bias, I describe which mechanisms drive the heterogeneous race match effects on learning skills and behavior.

3 Empirical Strategy

3.1 Los Angeles Unified School District Administrative Data

This study relies on an administrative panel dataset for students in grades 6 through 12 from LAUSD spanning the academic years of 2006-07 through 2016-17 ($N = 2,556,132$). As the second largest school district in the U.S., LAUSD is a unique setting to study race match effects on noncognitive skills for several reasons. First, the racial distributions of teachers and students within LAUSD allows for insight into traditionally understudied Latino and Asian populations. Second, LAUSD is located in an urban setting which tends to have higher shares of students of color than non-urban settings. Third, beyond the traditional measures of suspensions and absences, teachers within this district assess students based on other key noncognitive skills: cooperation and work habits.

For each student, LAUSD records race,[2] grade, gender, English Learner (EL) status, free- or reduced-priced lunch (FRL) eligibility, and student with disability (SWD) status. The data also link students to teachers by year, as well as their characteristics such as race, gender, years of experience,[3] and certification status (e.g., preliminary or fully certified). Using these data, I am able to generate a race match variable which indicates if a student and teacher share the same race. Since students may have multiple teachers, and it is not obvious which teacher contributes most to improving a student's learning skills or behavior,

[2]Student and teacher race categories include: Asian, Black, Latino, and White. Filipino, Native American, Pacific Islander, and Multi-Race have been subsumed into an "Other" category due to lack of sample size.

[3]Years of experience is top-coded at 10 years.

I collapse the data to the student-year level. Table 1 summarizes each of the student characteristics, as well as the averaged teacher characteristics by student race. The majority of students of color are FRL eligible, a proxy measure for a student's socioeconomic status, with 84% of Latino students being eligible for free- or reduced-price lunches. As a reference, the average for all public school students in California is 59.4% (NCES, 2017). Also, White and Asian students, on average, have teachers with higher levels of experience and certification status than Black and Latino students. Since each of these characteristics is expected to influence a student's learning skills and behavior, they are controlled for in the equation in the Model section of this paper.

Each year, a student is assigned a mark for achievement, work habits, and cooperation for every course (e.g., math, science, social studies, etc.) they are enrolled. I use these marks to generate separate grade point averages for each category by averaging a student's marks across courses to the year level. For all students, the marks for achievement (*GPA*) are measured on a traditional 4-point scale.[4] The marks for work habits and cooperation are on a two-point scale with the values of 0 for "Unsatisfactory", 1 for "Satisfactory", and 2 for "Exceptional." Appendix F provides the rubric for how the marks for achievement, work habits, and cooperation are graded within LAUSD. Additional noncognitive outcomes included in the dataset are: days absent, days enrolled, times suspended, and days suspended for each student at the academic-year level. I use the number of days a student is absent and the number of days a student is enrolled to generate the percent of days a student is absent (*%Days Absent*). I also generate the average number of days a student is suspended per the number of times the student is suspended (*Avg. Days Susp*). The panel nature of the dataset allows me to identify whether a student is retained into the same grade the following year (*Held Back*). In the next section, I describe how these noncognitive measures differ across student race and race match status.

[4]Marks for achievement are converted from letter grades to numbers using the following scale: "A" = 4, "B" = 3, "C" = 2, "E" = 1, and "F" = 0.

3.2 Descriptive Findings

As seen in Figure 1, LAUSD has a larger share of Latino teachers and students than average in the U.S. Nationally, Latino teachers make up 7.9% of the teaching workforce, while almost a quarter (24.7%) of the student population are Latino. In LAUSD, for students in grades 6 through 12 during the academic years of 2007-08 through 2017-18, these percentages are much larger than the national average at 27.1% and 75.8% for teachers and students, respectively. Figure 2 illustrates the differences in the treatment variable (*Share RM*) across student race in LAUSD. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year.[5] On average, students are race matched to 32% of their teachers in a given year.[6] Asian, Black, and Latino students are race matched to 14, 26, and 32 percent of their teachers, respectively. Despite making up a relatively smaller share of the student composition, White students are race matched to 61% of their teachers. A contributing factor to this phenomenon is that White students tend to attend schools taught by mostly White teachers (NCES, 2020). In addition, White teachers also make up the plurality at 44.6% in LAUSD. While both of these factors increase the probability that a White student will be placed into a race matched classroom, the latter only applies to White students, thus giving them an inequitable opportunity to receive the potential benefits of race matched instruction.

Figure 3 summarizes the averages of different student outcomes by student race and illustrates the gaps found in various noncognitive skills and outcomes between Asian/White students and Black/Latino students. On average, Asian and White students have higher GPAs, marks for work habits, and marks for cooperation than Black and Latino students. Asian and White students are also less likely to be held back, suspended, and miss fewer days. These gaps in noncogitive skills are consistent with the existing literature and illustrate the need to better serve students of color. Table 2 summarizes each noncognitive skill and outcome by race match status.[7] To account for the continuous nature of the treatment variable,

[5]Presumably, the amount of time a student and teacher are together, the larger impact that teacher may have on a student's learning skills and behavior. To account for this, *Share RM* is weighted by the number of times a teacher and student share a course/subject. That is, $ShareRM_{it} = \frac{\sum RM_{ict}}{Tch_{ict}}$ where RM_{ict} is an indicator for whether a student and teacher share the same race in a given course and year and Tch_{ict} is the total number of teacher-course combinations a student has in that same year.

[6]See Table 1 for the complete table of student summary statistics.

[7]See Table A1 in the Appendix A to view a version of Table 2 that is not separated by $\%RM$.

I separate out the descriptive statistics by students that are race matched to the majority of their teachers ($> 50\%\ RM$, "race matched") and those that are not ($< 50\%\ RM$, "non-race matched"). On average, "race matched" students share the race of 65.9% of their teachers while "non-race matched" students share the race of 18.6% of their teachers. A pattern that emerges is that, without controlling for confounding factors, race matched students tend to have similar or better noncognitive outcomes than their non-race matched peers. A notable exception is that race matched Black students have lower scores in learning skills and worse measures of behavior than non-race matched Black students. However, it is important to point out that these descriptive statistics do not control for any characteristics that are expected to impact student outcomes. For example, schools with higher concentrations of low-income students tend to also receive fewer resources for instruction which could lead to worse outcomes for these students (Darling-Hammond, 1998). Thus, to accurately measure a race match effect, I control for factors correlated with race matching and noncognitive outcomes, as explained the upcoming Model section of this paper.

3.3 Latent Learning Skills and Behavior

A student's noncognitive ability is difficult to directly measure. I use a latent variable framework and define two aspects of noncognitive ability. Specifically, "learning skills" measures a student's ability to use prior knowledge to accomplish tasks, evaluate their own work, and cooperate with others. On the other hand, "behavior" is a measure of a student's externalizing behaviors such as aggression, disruptiveness, or impulsiveness. While a student's learning skills or behavior are not directly measured, proxies of these variables may be generated using the information provided from observable student outcomes.

In this study, I categorize each noncognitive outcome as a "learning skill" or a "behavior" component. However, there is expected overlap across these two latent variables. For example, teachers assign value to a student's behavior, attitude, work habits, and effort when assigning grades (Guskey, 2018; Howley et al., 2000). Since a student's behavior and learning skills are linked to influencing GPA, this component could easily be used to proxy for either latent variable. However, in the specific context of this paper, the rubric that teachers use in LAUSD does not contain information on externalizing behaviors when assigning grades to students. For this reason, I have excluded GPA from entering the student's behavior function. Similarly, I include a student's work habits and cooperation marks into measuring their learning skills but not behavior.

Duckworth et al. (2007) find that students with fewer absences, suspensions, and grade retention are associated with higher levels of conscientiousness. Since low levels of conscientiousness have been linked to higher levels of externalizing behaviors (Barbbaranelli et al., 2003), I include these components into calculating the proxy for a student's behavior.

Several of the noncognitive skills and outcomes may also be influenced by a student's cognitive ability. Continuing with GPA as an example, teachers are tasked with measuring a student's "quality of work" by determining if the student "demonstrates an exemplary level of understanding" which could certainly be interpreted as a cognitive skill. In an attempt to parse out as much information on noncognitive skills as possible, the next section describes how confirmatory factor analysis uses the shared variation of these components to generate proxies for the learning skills and behavior measures.

3.3.1 Confirmatory Factor Analysis

Confirmatory factor analysis utilizes the shared variation of each pre-selected component (e.g., GPA, work habits, suspensions) in order to generate the proxies for each latent noncognitive variable: learning skills and behavior. More generally, for a given p noncognitive components, let $X = LF + U$, where X is a $p \times n$ matrix of p components for each of the n observations, F is a $k \times n$ matrix of $k < p$ factors, L is a $p \times k$ factor loading matrix, and U is a $p \times n$ uniqueness matrix which captures the variation unique to each component that is not shared by the other components. L is estimated via eigenvector decomposition by solving $Cov(X) = Cov(LF + U)$. Finally, I employ the Bartlett scoring method to calculate each proxy. That is, $F = (\hat{L}'\hat{\Psi}^{-1}\hat{L})^{-1}\hat{L}'\hat{\Psi}^{-1}$, where $\hat{\Psi}$ is a diagonal matrix whose elements are equal to the variance of the unique factor scores (Bartlett, 1937). The exact Bartlett scores for each component, along with a description, of the learning skills and behavior proxies are provided in the following sections.

3.3.2 Learning Skills Proxy

Using confirmatory factory analysis, I generate a learning skills proxy that indexes a student's ability to use prior knowledge to accomplish tasks, evaluate their own work, and cooperate with others. Each noncognitive component (i.e., GPA, marks for work habits, and marks for cooperation) used in the process detailed above contains information about a child's learning skills. Confirmatory factor analysis attempts to parse out this information from each of the components to define the latent learning skills variable by utilizing the shared or common variation across the three components used, while ignoring the variation unique to each component. The resulting equation for the learning skill measure is:

$$LearningSkills = 0.29 * GPA + 0.54 * WorkHabits + 0.20 * Cooperation. \tag{1}$$

This measure is standardized across the sample to generate a *learning skills index* (LSI) with higher values being associated with “better" learning skills. As a reference, Petek and Pope (2021) find that a 1sd increase in a similar measure is expected to decrease the probability that a student drops out of school by 4 percentage points. The correlations between each noncognitive component can be found in Table A2 in the Appendix. Table 3 displays the factor loadings and Eigenvalue for the learning skill index. The factor loadings can range from [-1,1] and represent the correlation between the component and the proxy (or factor) variable. While related, these values are different from the Bartlett scores used to create the proxy. For the learning skills index, each component is highly and positively correlated to the proxy with each factor loading being above 0.95. As a rule of thumb, having an Eigenvalue greater than one indicates that the proxy is a good measure. In the case for the learning skills proxy, the Eigenvalue is 2.78 and indicates that the shared information across the components used explain 93% of the total variation.

3.3.3 Behavior Proxy

I also use confirmatory factor analysis to create a proxy for a student's behavior. This proxy captures a student's externalizing behaviors such as aggression, disruptiveness, or impulsiveness. The Bartlett scores associated with each noncognitive component of the behavior proxy are shown in the following equation:

$$Behavior = -0.73 * \%DaysAbsent - 0.24 * Suspensions - 0.50 * GradeRetention. \quad (2)$$

As with the learning skills index, the proxy for behavior is standardized across observations with higher scores being associated with better behavior. I refer to the standardized behavior measure as the *behavior index* (BI). For reference, Jackson (2018) indicates that an increase of 1sd in a similar measure is estimated to increase the probability a student graduates from high school by 15.8 percentage points. The factor loadings and Eigenvalues for the behavior index is shown in Table 3. Each component is negatively correlated with the behavior proxy, and the Eigenvalue is greater than one, the rule of thumb threshold for a good proxy measure.

Panel (f) of Figure 3 illustrates the averages of the learning skills index and behavior index by student race. The gap for each noncognitive skill index between Asian and White students and Black and Latino students is made apparent. The indices are standardized across the sample and are interpreted as standard deviations from the mean. As seen in this figure, on average, Asian and White students have learning skills and behavior measures well above zero, while these measures for Black and Latino students are below the mean. Since noncognitive skills impact long-run outcomes, it is vital to shrink these gaps which may occur through student-teacher race matching. Table 2 summarizes the learning skills index and behavior index by race match status. Descriptively, students in race matched classes tend to have higher scores in both indices. The exception is that race match Black students have a lower learning skills index and behavior index score than non-race matched students from a descriptive standpoint.

3.4 Model & Identification Strategy

A potential threat to identifying race match effects is that students may sort by unobservable characteristics to their teachers based on race, which would bias the estimate. Since students and teachers sort by race across schools (NCES, 2020), I include school-grade fixed effects to eliminate this concern by identifying the race match effect using the variation found within a school-grade. I also include student fixed effects to control for any time invariant unobserved student characteristics that may lead to sorting. While these two fixed effects are used in the main specification, I also analyze other models in the Robustness section to further investigate the concern for student and teacher sorting. I estimate the impact that student-teacher race matching has on learning skills and behavioral measures by using the following model:

$$Y_{igst} = \beta_0 + \beta_{RM}(ShareRM_{igst}) + \mathbf{X_{igst}}\mathbf{\Gamma_1} + \mathbf{\bar{Z}_{igst}}\mathbf{\Gamma_2} + \mathbf{\bar{X}_{gst}}\mathbf{\Gamma_3} + \theta_i + \theta_{gs} + \theta_t + \epsilon_{igst}, \tag{3}$$

where Y_{igst} is the outcome of interest for student i in grade g in school s at year t. $ShareRM$ is the share of race matched teachers for the student, $\mathbf{X}$ is a vector of time-variant student characteristics, $\mathbf{\bar{Z}}$ is a vector of averaged teacher characteristics, $\mathbf{\bar{X}}$ is a vector of classroom characteristics, θ_i, θ_{gs}, and θ_t are student, school-grade, and year fixed effects, respectively. ϵ is a normally distributed error term. I repeat this analysis by student race using:

$$Y_{igst} = \beta_0 + \beta_{RM}(\mathbb{1}(Race_i) * ShareRM_{igst}) + \mathbf{X_{igst}}\mathbf{\Gamma_1} + \mathbf{\bar{Z}_{igst}}\mathbf{\Gamma_2} + \mathbf{\bar{X}_{gst}}\mathbf{\Gamma_3} + \theta_i + \theta_{gs} + \theta_t + \epsilon_{igst}, \tag{4}$$

where each component is identical to equation 3 with the exception that $ShareRM$ is now multiplied by an indicator variable, $\mathbb{1}(Race)$, which represents the student's race. The standard errors are clustered at the student level, which is the unit of treatment in this study.

Even though a student's outcome in a given year is endogenous with respect to teachers, the race match effect is identified from the within student variation in the share of race matched teachers a student has across years. That is, if a particular teacher favorably grades a race matched student within a given year, this bias would average out as long as the student remains in LAUSD for a sufficient amount of years. However, the race match effect could still be biased if all teachers of a particular race favorably grade race match students across all years. In this case, the race match effect would no longer be interpreted as a change in the student's noncognitive skills, but rather a change in the grading of the student's noncognitive skills. Students of color could still benefit under this interpretation if the student's view more favorable grading as affirmation from their teacher. In a meta-analysis, Wu et al. (2021) provide examples of several studies which find that teacher affirmation improves outcomes, especially for disadvantaged groups (e.g, students of color, female, English learners).

3.5 Separating student and teacher race match mechanisms

The subjective nature of noncognitive skills make them difficult to study in a causal way. Recently, Elder and Zhou (2021) developed a method which "ties down" these subjective measures to more objective measures, such as test scores, in order to control for biases that teachers may inflict onto marks for noncognitive skills. Intuitively, a teacher's assessment of a student's development in learning skills may be influenced by the other students in the classroom. A child that seems to exhibit low effort when surrounded by high effort students may appear high effort when surrounded by low effort students even when the child does not change their actual level of effort. Elder and Zhou's approach removes this *reference bias* by assuming that the underlying bias for an objective measure is similar to a subjective one and adjusting the subjective outcome by the amount of bias found in the objective measure. I make a slight modification to the authors' methods to also include a student's race when calculating the reference bias. I use math test scores as the baseline measure when calculating the reference biases, which means that my teacher reference bias analyses necessarily restricts the sample to only students with test scores. While I am still unable to separate out passive and active channels, this approach controls for teacher bias which leaves the race match effect that is driven by the student mechanisms described in Section 2. In other words, if the effect with the adjusted outcomes is smaller than the original effect, then the original race match effect must have been driven by teacher-side mechanisms. Similarly, if the race match effect with the adjusted outcomes is larger or equal to the original effect size, then the main driver in the race match effect is a student-side mechanism.

Elder and Zhou (2021) propose two separate methods, each requiring a different assumption, to adjust subjective outcomes for reference bias. I slightly modify and employ both of these methods and make adjustments to the GPA, marks for work habits, marks for cooperation, suspensions, and grade retention. Then, using these newly adjusted measures, I recalculate each version of the learning skills and behavior indices. Finally, I reanalyze equation 4 using the reference bias adjusted learning skills and behavior indices. I provide a detailed explanation on calculating the reference bias adjusted measures for condition 1 and condition 2 in Appendix B.

3.5.1 Condition 1: Similar ratio of within-school-race and between-school-race variance across skills

In the first reference bias adjustment, I assume that the each subjective and objective measures are similarly distributed within a school-grade-year-race (SGYR) level. Elder and Zhou (2021) show that under this assumption, the ratio of the within-SGYR variation of any (subjective or objective) measure to the between-SGYR variation will be constant. I calculate this ratio using math test scores as the baseline and find the measure-specific reference bias for each learning skill component by using the above assumption and the ratio of within-/between-SGYR variation in math test scores. Finally, I adjust the noncognitive skill component for each student within the SGYR level by subtracting the calculated reference bias and standardizing the newly adjusted outcome to have the same variance as the original measure.

3.5.2 Condition 2: Teacher's bias does not vary across skills

The second reference bias adjustment is similar to the first except, under the assumption that all skills at the SGYR level are biased by the same amount, there only needs to exist an exemplar noncognitive measure to calculate the reference bias. Using GPA, I repeat the steps under condition 1, but use the resulting reference bias to adjust the remaining noncognitive components. Table B1 shows the adjusted outcomes for condition 1 and condition 2 for each noncognitive measure by race and race match status.

For example, Asian students matching less than half of their teachers' races have an unadjusted GPA of 3.222. The adjusted GPAs for these students is 3.089 and 3.222 under conditions 1 and 2, respectively.

4 Race Match Effects on Noncognitive Skills and Outcomes

4.1 Individual Noncognitive Skills and Outcomes

Table 4 shows the effects of student-teacher race matching on the individual noncognitive measures. Each learning skills component (e.g., GPA, marks for effort, and marks for cooperation) are standardized at the grade-year level, and the coefficients for these outcomes are interpreted as a change in standard deviations. Since the effects found in this table for *Share RM* are interpreted as a student going from no race matched teachers to all race matched teachers, I interpret the results as the effect of a 10 percentage point increase in *Share RM* by dividing each coefficient by ten. Panel A shows that the overall race match effect for GPA, marks for work habits, and marks for cooperation are positive with a 10 percentage point increase in the share of race matched teachers leading to an expected increase of 0.0029sd, 0.0037sd, and 0.0024sd for each component, respectively. Panel B shows how the race match effects differ across race with Asian, Black, and Latino students expected to experience increases in the learning skills with an increase in the share of race matched teachers. However, White students do not seem to benefit from having a larger share of race matched teachers. In fact, a 10 percentage point increase in the share of White teachers is expected to decrease marks for cooperation by 0.0019sd for White students. This is a key result because these findings indicate that increasing the share of teachers of color will not detract from the development of learning skills for White students. To better understand the magnitude of these effect sizes across student race, Figure 4 plots each coefficient with a 95-percent confidence interval. The race match effect on each learning skills component is largest for Asian and Black student but also remains significantly positive for Latino students.

Held Back and *Suspension* in Table 4 are indicator variables and the effects are interpreted as a change in the probability of the outcome occurring. Figure 5 illustrates the coefficient plots and confidence bands for these two variables. The overall effects indicate that an increase in the share of race matched teachers has no compelling impact on either variable. While the effect found on *Held Back* is statistically significant, a student going from zero race matches teachers to all race match teachers would be expected to be retained in their grade by 0.3 percentage points. These effects remain insignificant even when examining them by race. For example, Asian students are expected to find that a 10 percentage point increase in the share of race matched teachers decreases their probability of being held back by 0.1 percentage points or suspended by 0.06 percentage points. While these effect sizes are negligible, the information these variables contain on noncognitive skills still contribute to the race match effect found for the behavior index in the next section. Note that these results, as with all of the results in this paper, do not take into account the possibility for the race match effect to compound over time.

For column (6) of Table 4, I use a log transformation for the number of days the student is absent (*Days Absent*) and interpret the results as the percent change in the number of days absent.[8] The coefficient plots can be found in Figure 6. The overall race match effect indicates that a 10 percentage point increase in the share of race matched teachers is expected to decrease the number of days a student is absent by 0.25 percent. When looking by race, increasing the share of race matched teachers by 10 for Asian and Black students is expected to decrease the number of days missed by 0.99 and 0.67 percent, respectively. Again, these values may seem small, but they do not account for the possibility of a compounding effect as a student is race matched multiple times throughout their educational career.

4.2 Learning Skills and Behavior Indices

The race match effect for learning skills and behavior are found in Table 5. Panel A shows the results for equation 3 which estimates the overall race match effect regardless of a student's race. I find that increases in the share of race matched teachers lead to a higher learning skills. Column (1) indicates that a 10 percentage point increase in the share of race matched teachers is associated with a 0.0011 sd increase in the learning skills index. Using the results from Petek and Pope (2021), this estimate

[8]To account for zero values, I use a $ln(y + 1)$ transformation.

suggests a decrease in the probability a student drops out of high school by 0.0045 percentage points. Panel B displays the results for equation 4 which looks at the heterogeneous race match effect by student race. For Asian, Black, and Latino students, a 10 percentage point increase in the share of race matched teachers is estimated to improve the learning skills by 0.0137 sd, 0.0096 sd, and 0.0038 sd, respectively. These improvements to learning skills are estimated to decrease the probability Asian students drop out of school by 0.056 and Black students by 0.016 percentage points with Latino students in-between.

Panel A of Table 5 also indicates that student behavior also benefits from race matching. A 10 percentage point increase in the share of race matched teachers is expected to increase a student's behavior index by 0.0041sd. The results from Jackson (2018) implies an increase in the probability of graduating high school by 0.055 percentage points for these students. Asian and Black students seem to benefit the most from race matching in terms of improvements to their behavior index with a 10 percentage point increase the in share of race matched teachers leading to an expected decrease in the probability of dropping out of high school by 0.067 percentage points.

Figure 7 compares the race match effect on the learning skills index to the effect on the behavior index across each race. In general, the effect on learning skills seems larger than that on behavior. However, there is a significant, positive impact of race matching on behavior, especially for Asian and Black students.

4.2.1 Reference Bias Results and Potential Mechanisms

Table 6 shows the results of the reference bias adjusted learning skills index and behavior index. While the estimates make no distinction between active or passive channels, they do account for a teacher's bias when assigning students a mark for achievement, effort, cooperation, suspension, or grade retention. This effectively teases out the race match mechanisms driven by teachers and leaves student mechanisms, such as the role model effect (active channel) or stereotype threat (passive channel). By comparing the original specification results to the those adjusted for reference bias, I can infer which mechanism is the main driver for the race match effect.

Both reference bias adjustment conditions rely on an objective measure to act as a baseline in calculating a teacher's bias. Since I use a student's math score, this restricts the sample to those students with a math test score. I repeat the analysis on the learning skills and behavior following equation 4 for this subsample. As seen in Figure 8, the results for the restricted sample are similar in magnitude to the results for the main sample for both indices. The race match effect on learning skills for Asian and Black students are similar when controlling for teacher bias. This indicates that these effects are mainly driven by student-side mechanisms. It is possible that race matched Asian and Black students benefit from a lack of stereotype threat. As explained in a previous section, an awareness of any negative racial stereotype could cause a student to experience anxiety when not race matched to their teacher which may cause them act in ways that are perceived as misbehaving. This anxiety could also prevent them from participating in classrooms which, according to Guskey (2018), is a factor teachers use when assigning grade points to students. A student's level of anxiety would be reduced if they were to not experience racial stereotype threats in race matched classrooms which, in turn, could improve behavior and allow for an uninhibited development of learning skills. Panel (b) of Figure 8 indicates that the race match effect on learning skills for Latino students decreases when accounting for a teacher's reference bias. This indicates that a teacher-side mechanism is the driving force behind the race match effect. Since LAUSD has one of the largest populations of Latino teachers in the country, it is possible that Latino students benefit from culturally relevant pedagogy which improves a student's connection to the course materials (Gay, 2000). While the coefficients indicate the race match effect on behavior skills for Asian and Black students are driven by student mechanisms, Figure 8 reveals that these measures are not statistically different from one another, leaving ambiguity on exactly which mechanisms are driving the results.

5 Robustness Checks

A threat to identifying the race match effect is that students and teachers may sort by race for reasons that influence their behavior or learning skills. I address this issue in equation 3 by including student fixed effects, which utilizes within-year variation and eliminates the concern of any time invariant student characteristic that causes a student to sort, and school-grade fixed effects, which mitigates the concern of students and teachers sorting across schools. However, students that sort in one year may also be more likely to sort in a following year. As further checks for sorting, I also analyze two additional specifications. First, I analyze and compare different levels of fixed effects by modifying equation 4. Changing the level of fixed effects alters the necessary identifying assumptions and finding similar results between the different models mitigates the concern for sorting bias. Second, I include lagged outcomes in the regression model, which control for the case of students with similar levels of perceived behavior or learning skills being sorted into the same classrooms. I do not find evidence of sorting by race in either robustness check when taking into account student, teacher, and school characteristics, which indicates that the method used in equations 3 and 4 is sufficient to account for sorting bias.

5.1 Comparing results of specifications with different levels of fixed effects

Harbatkin (2021) checks for student-teacher sorting by comparing the effects from several modified regression models. By changing the levels of fixed effects used, they compare within-year, across course estimates to across year, within-course estimates. The race match effect is identified when exploiting within-year, across course variation if students that are likely to sort into race matched classrooms by year do not do so by subject. Similarly, if students are likely to sort by subject, but not by year, then the model exploiting the across year, within-course variation would be identified. If comparing the effects of these two models results in similar effect sizes, then it is likely that these results are an unbiased estimate of the race match effect. Of course, if there are unobserved and time-variant student characteristics that are correlated with the outcome and these students are sorting within-year and within-course, these methods would still be biased. I repeat the exercise from Harbatkin (2021) using the students in grades 6 through 12. Since students in earlier grades tend to only have a single teacher, there would not be enough variation in the race match status to estimate the within-year, across course effect. For a similar reason, I am only able to look at the effects of GPA, effort, and cooperation since the behavior outcomes are collected at the year level. The models I use are similar to equation 4 with two key differences: 1) course data is added to the model and is represented by c and 2) the fixed effect levels vary. The equations are given by:

$$Y_{icgst} = \alpha + \beta_{RM}(\mathbb{1}(Race_i) * ShareRM_{igst}) + \mathbf{X_{icgst}}\mathbf{\Gamma_1} + \mathbf{Z_{icgst}}\mathbf{\Gamma_2} + \mathbf{\bar{X}_{cgst}}\mathbf{\Gamma_3} + \theta_{gs} + \theta_i + \theta_t + \theta_c + \epsilon_{igst}, \tag{5}$$

$$Y_{icgst} = \alpha + \beta_{RM}(\mathbb{1}(Race_i) * ShareRM_{igst}) + \mathbf{X_{icgst}}\mathbf{\Gamma_1} + \mathbf{Z_{icgst}}\mathbf{\Gamma_2} + \mathbf{\bar{X}_{cgst}}\mathbf{\Gamma_3} + \theta_{gs} + \theta_{ic} + \theta_t + \epsilon_{igst}, \tag{6}$$

$$Y_{icgst} = \alpha + \beta_{RM}(\mathbb{1}(Race_i) * ShareRM_{igst}) + \mathbf{X_{icgst}}\mathbf{\Gamma_1} + \mathbf{Z_{icgst}}\mathbf{\Gamma_2} + \mathbf{\bar{X}_{cgst}}\mathbf{\Gamma_3} + \theta_{gs} + \theta_{it} + \theta_c + \epsilon_{igst}, \tag{7}$$

where equation 5 includes separate fixed effects for students, year, and courses, equation 6 identifies off of across year, within-course variation, and equation 7 identifies off of within-year, across course variation. The standard levels for each model are clustered at the student, student-course, and student-year level, respectively. Tables C1 and C2 show the results for all three methods of fixed effects and reveal similar effects within each noncognitive skill measure: GPA, work habits, cooperation, and the learning skills index. This provides some evidence that the results are robust to student-teacher sorting. Since the analysis in this section uses course-level data, the effect is measuring a direct race match effect between student and teacher. This is different from the results in the main analysis which is interpreted as the effect of the share of race matched teachers. As such, comparisons of the effects from the course-level analysis to the main specification should be made with this understanding in the differences in interpretation. With this in mind, Figures 9 through 12 plot the coefficients from the above models to the effect found in the main specification. Overall, the race match effects for Asian, Black, and Latino students remains positive between the course-level and main specifications. However, for Latino students, the direct race match effect on the learning skills components seem to have a larger effect than the estimates using the share of race matched teachers indicate.

5.2 Lagged covariates and outcomes in the model

Including a lag of the outcome variable of interest has been used as another way to control for the potential problem of student-teacher sorting (Kane and Staiger, 2008; Kane et al., 2013). Chetty et al. (2014) use a student's prior test scores when calculating value-added measures, and Liu and Loeb (2021) control for prior absence rates when studying a teacher's impact on student attendance. I include a vector of lagged measures of GPA, marks for effort, marks for cooperation (where applicable), days absent, and times suspended to equation 4:

$$Y_{igst} = \beta_0 + \beta_1 Y_{igst-1} + \beta_{RM}(ShareRM_{igst}) + \mathbf{X_{igst}\Gamma_1} + \mathbf{\bar{Z}_{igst}\Gamma_2} + \mathbf{\bar{X}_{gst}\Gamma_3} + \theta_i + \theta_{gs} + \theta_t + \epsilon_{igst}. \quad (8)$$

The results can be found in Table C3. Figures 13 through 17 compares the race match effects between the main specification and equation 8 for each noncognitive skill and measure. Overall, these effects are robust to either specification.

5.3 Exploratory Factor Analysis

To capture noncognitive information from additional variables and to separate any cognitive measures being captured by the learning skills and behavior indices, I also perform an exploratory factor analysis[9] which, in addition to the original noncognitive skill components, includes math and English Language Arts (ELA) test scores. However, test scores only exist for students in grades 3 through 8 and grade 11, which restricts the sample to those students with test scores.[10]

Table 7 provides the factor loadings as well as the Eigenvalues for each index. I find that the learning skills components (i.e., GPA, work habits, and cooperation) have the highest factor loadings onto the first factor and deem this to be the learning skills index. Similarly, test scores load the highest onto the second factor and appear to measure cognitive skills. Importantly, in exploratory factor analysis, each factor is

[9]Exploratory factor analysis follows the steps found in section 3.3.1 with the objective to extract the maximum shared variance possible from the components regardless of the number of factors needed to do so.

[10]In 2014, California changed testing administrations from the California State Test to the Smarter Balanced Assessment System. Tests scores for the pilot year were not released resulting in no test scores in the data for the 2013-2014 academic year.

orthogonal to one another. In other words, the learning skills index in column (1) is uncorrelated with the cognitive skills index in column (2). The behavioral measures load the most onto the third factor, but the Eigenvalue is less than one, which indicates this measure does not explain enough of the total variation to be a useful factor. The final column represents the uniqueness or residual of each noncognitive and test score component.

Table 8 and Figure 18 compare the main race match results of the learning skills index created using confirmatory factory analysis to the index generated via exploratory factor analysis. The overall race match effect is larger under the exploratory factor analysis version of the learning skills index, which indicates that race matched teachers have more impact on the noncognitive aspects of learning skills than than the cognitive aspects. When looking by race, it becomes clear that this effect is mainly driven by the effect found for Latino students.

6 Conclusion & Implications

This paper adds to the growing literature which shows that students of color benefit from matching the race of their teacher. The mounting empirical evidence alludes to student-teacher race matching as a potential mechanism for shrinking the educational gaps between White and non-White students. While most of the existing research has focused on cognitive measures, such as test scores, a growing body of literature also points to race matching to improve a student's noncognitive skills. Developing noncognitive skills is essential because these measures improve long-run outcomes such as college enrollment, employment, and wages.

I find that students of color in grades 6 through 12 seem to benefit the most from increased race matches with improvements to learning skills and behavior, decreases in grade retention, and fewer missed days of school. When students are placed with a teacher of the same race, they may be less likely to worry about exhibiting negative racial stereotypes. Evidence points to stereotype threat, a student mechanism, being a potential factor in the race match effect on learning skills for Asian and Black students since the results are persistent even when accounting for teacher bias. However, teacher mechanisms seem to be the driving factor for the race match effect on a student's learning skills for Latino students. Since LAUSD has a large share of Latino teachers and students, a plausible explanation could be that teachers are able to administer culturally relevant pedagogy.

A significant policy implication of this study is to further diversity the teacher workforce. While the racial composition of students continues to diversify, White teachers remain the vast majority. In fact, over 70 percent of education majors are White (Office of Planning, Evaluation and Policy Development, 2016). Having more teachers of color in the labor supply would ensure higher probabilities for students of color to benefit from being in race matched classrooms. This goal can by achieved by implementing policies to attract and train more people of color into teaching.

References

C. Barbbaranelli, G. Caprara, A. Rabasca, and C. Pastorelli. A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences*, 34:645–664, 2003.

M. S. Bartlett. The statistical conception of mental factors. *British Journal of Psychology. General Section*, 28(1):97–104, 1937. doi: https://doi.org/10.1111/j.2044-8295.1937.tb00863.x.

H.-Y. S. Cherng and P. F. Halpin. The importance of minority teachers: Student perceptions of minority versus white teachers. *Educational Researcher*, 45(7):407–420, 2016. doi: 10.3102/0013189X16671718. URL https://doi.org/10.3102/0013189X16671718.

R. Chetty, J. Friedman, and J. Rockoff. The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679, 2014.

M. J. Chin, D. M. Quinn, T. K. Dhaliwal, and V. S. Lovison. Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, 49(8):566–578, 2020. doi: https://doi.org/10.3102/0013189X20937240.

L. Darling-Hammond. Unequal opportunity: Race and education, 1998. URL https://www.brookings.edu/articles/unequal-opportunity-race-and-education/.

T. S. Dee. Teachers, Race, and Student Achievement in a Randomized Experiment. *The Review of Economics and Statistics*, 86(1):195–210, 2004.

T. S. Dee. Social Identity and Achievement Gaps: Evidence From an Affirmation Intervention. *Journal of Research on Educational Effectiveness*, 8(2):149–168, 2015. doi: 10.1080/19345747.2014.906009.

D. Downey and S. Pribesh. When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education*, 77, 01 2004. doi: 10.1177/003804070407700401.

A. L. Duckworth, C. Peterson, M. D. Matthews, and D. R. Kelly. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6):1087–1101, 2007. doi: 10.1037/0022-3514.92.6.1087.

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

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

T. Elder and Y. Zhou. The black-white gap in noncognitive skills among elementary school children. *American Economic Journal: Applied Economics*, 13(1):105–32, January 2021. doi: 10.1257/app.20180732. URL https://www.aeaweb.org/articles?id=10.1257/app.20180732.

R. F. Ferguson. Teachers' perceptions and expectations and the black-white test score gap. *Urban Education*, 38(4):460–507, 2003. doi: 10.1177/0042085903038004006. URL https://doi.org/10.1177/0042085903038004006.

L. Fox. Seeing potential: The effects of student–teacher demographic congruence on teacher expectations and recommendations. *AERA open*, 2(1):2332858415623758, 2015.

J. Fryer, Roland G. and S. D. Levitt. Testing for racial differences in the mental ability of young children. *American Economic Review*, 103(2):981–1005, April 2013. doi: 10.1257/aer.103.2.981. URL https://www.aeaweb.org/articles?id=10.1257/aer.103.2.981.

G. Gay. *Culturally Responsive Teaching: Theory, Research, and Practice*. 01 2000.

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

S. Gershenson, C. M. D. Hart, J. Hymna, C. A. Lindsay, and N. W. Papageorge. The Long-Run Impact of Same-Race Teachers. *National Bureau of Economic Research*, (Working Paper 25254), 2018.

J. A. Grissom, E. C. Kern, and L. A. Rodriguez. The "Representative Bureaucracy" in Education: Educator Workforce Diversity, Policy Outputs, and Outcomes for Disadvantaged Students. *Educational Researcher*, 44(3):185–192, 2015. doi: 10.3102/0013189X15580102.

J. A. Grissom, L. Rodriguez, and E. Kern. Teacher and Principal Diversity and the Representation of Students of Color in Gifted Programs. *The Elementary School Journal*, 117(3):396–422, 2017.

L. J. Guskey, Thomas R.; Link. Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy & Practice*, 26(3):303–320, 2018. ISSN 0969-594X. doi: 10.1080/0969594X.2018.1555515. URL `https://browzine.com/articles/253609902`.

J. Harackiewicz, C. Rozek, C. Hulleman, and J. Hyde. Helping parents to motivate adolescents in mathematics and science: An experimental test of a utility-value intervention. *Psychological science*, 23:899–906, 07 2012. doi: 10.1177/0956797611435530.

E. Harbatkin. Does student-teacher race match affect course grades? *Economics of Education Review*, 81:102081, 2021. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2021.102081. URL `https://www.sciencedirect.com/science/article/pii/S0272775721000042`.

A. Hashim, K. Strunk, and T. Dhaliwal. Justice for all? suspension bans and restorative justice programs in the los angeles unified school district. *Peabody Journal of Education*, 93, 02 2018. doi: 10.1080/0161956X.2018.1435040.

J. Heckman, J. Stixrud, and S. Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482, 2012.

J. Heckman, R. Pinto, and P. Savelyev. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–86, October 2013. doi: 10.1257/aer.103.6.2052. URL `http://www.aeaweb.org/articles?id=10.1257/aer.103.6.2052`.

A. Howley, P. S. Kusimo, and L. Parrott. Grading and the ethos of effort. *Learning Environments Research*, 3(3):229–246, 2000. doi: https://doi.org/10.1023/A:1011469327430.

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

J. L. Jennings and T. A. DiPrete. Teacher effects on social and behavioral skills in early elementary school. *Sociology of Education*, 83(2):135–159, 2010. doi: 10.1177/0038040710368011. URL `https://doi.org/10.1177/0038040710368011`.

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

T. J. Kane and D. O. Staiger. Estimating teacher impacts on student achievement: An experimental evaluation. *National Bureau of Economic Research*, December(Working Paper 14607), 2008.

T. J. Kane, D. F. Mccaffrey, T. Miller, and D. O. Staiger. Have we identified effective teachers? validating measures of effective teaching using random assignment. In *Research Paper. MET Project. Bill & Melinda Gates Foundation*, 2013.

G. Ladson-Billings. Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3):465–491, 1995. doi: 10.3102/00028312032003465. URL https://doi.org/10.3102/00028312032003465.

A. Lareau. *Unequal Childhoods: Class, Race, and Family Life*. 09 2011. ISBN 9780520949904. doi: 10.1525/9780520949904.

A. Lareau and E. M. Horvat. Moments of social inclusion and exclusion: Race, class and cultural capital in family-school relationships. *Sociology of Education*, 72(1):37–53, 1999.

H.-H. Lim. Representative bureaucracy: Rethinking substantive effects and active representation. *Public Administration Review*, 66(2):193–204, 2006. doi: https://doi.org/10.1111/j.1540-6210.2006.00572.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6210.2006.00572.x.

C. Lindsay and C. Hart. Exposure to Same-Race Teachers and Student Disciplinary Outcomes for Black Students in North Carolina. *Educational Evaluation and Policy Analysis*, 39(3):1–9, 2017.

J. Liu and S. Loeb. Engaging teachers: Measuring the impact of teachers on student attendance in secondary school. *Journal of Human Resources*, 56(2):N.PAG, 2021. ISSN 0022166X.

P. B. McGrady and J. R. Reynolds. Racial mismatch in the classroom: Beyond black-white differences. *Sociology of Education*, 86(1):3–17, 2013. doi: 10.1177/0038040712444857. URL https://doi.org/10.1177/0038040712444857.

H. Milner. Culturally Relevant Pedagogy in a Diverse Urban Classroom. *The Urban Review*, 43:66–89, 2011.

NCES. *Digest of Education Statistics: 2017* (NCES 2017- 070). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics., 2017.

NCES. *Statistical analysis report: Higher education* (NCES 97-584). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics., 2020.

G. L. S. C. Oates. Teacher-student racial congruence, teacher perceptions, and test performance*. *Social Science Quarterly*, 84(3):508–525, 2003. doi: https://doi.org/10.1111/1540-6237.8403002. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6237.8403002.

Office of Planning, Evaluation and Policy Development. U.S. Department of Education, Office of Planning, Evaluation and Policy Development, The State of Racial Diversity in the Educator Workforce., 2016.

J. Okonofua, J. P. Goyer, C. A. Lindsay, J. Haugabrook, and G. M. Walton. A scalable empathic-mindset intervention reduces group disparities in school suspensions. 2021. URL https://doi.org/10.31219/osf.io/z6kw2.

J. A. Okonofua, G. M. Walton, and J. L. Eberhardt. A vicious cycle: A social–psychological account of extreme racial disparities in school discipline. *Perspectives on Psychological Science*, 11(3):381–398, 2016. doi: 10.1177/1745691616635592. URL https://doi.org/10.1177/1745691616635592. PMID: 27217251.

A. Ouazad. Assessed by a Teacher Like Me: Race and Teacher Assessments. *Education Finance and Policy*, 9(3):334–372, 2014.

N. Petek and N. Pope. The Multidimensional Impact of Teachers on Students. *Working Paper*, 2021.

Pew Research Center. What Unites and Divides Urban, Suburban and Rural Communities, 2018. file:///D:/Downloads/Pew-Research-Center-Community-Type-Full-Report-FINAL.pdf.

J. Price. The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, 29(6):901 – 910, 2010.

D. D. Ready and D. L. Wright. Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, 48(2): 335–360, 2011.

S. F. Reardon and C. Galindo. The hispanic-white achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, 46(3):853–891, 2009. doi: 10.3102/0002831209333184. URL `https://doi.org/10.3102/0002831209333184`.

G. Rodriguez. From troublemakers to pobrecitos: Honoring the complexities of survivorship of latino youth in a suburban high school. *Journal of Latinos and Education*, pages 1–18, 2020. doi: 10.1080/15348431.2020.1796672. URL `https://doi.org/10.1080/15348431.2020.1796672`.

E. Saft and R. Pianta. Teachers' perceptions of their relationships with students: Effects of child age, gender, and ethnicity of teachers and childen. *School Psychology Quarterly*, 16(2):125–141, 2001.

M. Shirrell, T. Bristol, and T. Britton. The Effects of Student-Teacher Ethnoracial Matching on Exclusionary Discipline for Asian American, Black, and Latinx Students: Evidence From New York City. *Annenberg Working Paper*, October(EdWorkingPaper No. 21-475), 2021.

R. Skiba, R. Horner, C.-G. Chung, M. Rausch, S. May, and T. Tobin. Race is not neutral: A national investigation of african american and latino disproportionality in school discipline. *School Psychology Review*, 40, 03 2011. doi: 10.1080/02796015.2011.12087730.

C. M. Steele. Stereotype Threat and the Intellectual Test Performance of African Americans. *American Psychologist*, 69(5):613–629, 1995.

C. M. Steele and J. Aronson. A threat in the air: How stereotypes shape intellectual identity and performance. *Journal of Personality and Social Psychology*, 52(6):797–811, 1997.

H. R. Tenenbaum and M. D. Ruck. Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99(2):253–273, 2007.

P. Todd and K. Wolpin. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1):91–136, 2007. ISSN 19328575, 19328664. URL `http://www.jstor.org/stable/10.1086/526401`.

F. L. Van Voorhis, M. F. Maier, J. L. Epstein, and C. M. Lloyd. The impact of family involvement on the education of children aged 3 to 8. New York, MDRC, 2013.

A. Wright, M. A. Gottfried, and V. N. Le. A Kindergarten Teacher Like Me: The Role of Student-Teacher Race in Social-Emotional Development. *American Educational Research Journal*, 54(1S):78S–101S, 2017.

Z. Wu, T. F. Spreckelsen, and G. L. Cohen. A meta-analysis of the effect of values affirmation on academic achievement. *Journal of Social Issues*, 77(3):702–750, 2021. doi: https://doi.org/10.1111/josi.12415. URL `https://spssi.onlinelibrary.wiley.com/doi/abs/10.1111/josi.12415`.

Tables

Table 1: Student Descriptive Characteristics

	Asian	Black	Latino	White	Total
Female	0.482	0.501	0.490	0.474	0.489
	(0.500)	(0.500)	(0.500)	(0.499)	(0.500)
FRL	0.615	0.722	0.836	0.404	0.774
	(0.487)	(0.448)	(0.370)	(0.491)	(0.418)
EL	0.126	0.010	0.239	0.060	0.195
	(0.332)	(0.097)	(0.427)	(0.238)	(0.396)
SWD	0.030	0.067	0.048	0.046	0.048
	(0.172)	(0.251)	(0.213)	(0.209)	(0.214)
% Asian (Tch)	0.141	0.087	0.089	0.100	0.093
	(0.157)	(0.130)	(0.131)	(0.138)	(0.134)
% Black (Tch)	0.071	0.257	0.099	0.064	0.109
	(0.121)	(0.258)	(0.160)	(0.115)	(0.173)
% Latino (Tch)	0.198	0.174	0.320	0.176	0.287
	(0.195)	(0.182)	(0.245)	(0.177)	(0.239)
% White (Tch)	0.546	0.428	0.430	0.614	0.453
	(0.243)	(0.263)	(0.256)	(0.230)	(0.260)
Tot. Tch.	5.878	5.987	5.848	5.862	5.865
	(1.526)	(1.948)	(1.774)	(1.622)	(1.767)
% Female (Tch)	0.522	0.524	0.508	0.535	0.512
	(0.227)	(0.231)	(0.230)	(0.228)	(0.230)
% Cert. (Tch)	0.920	0.850	0.870	0.921	0.875
	(0.135)	(0.199)	(0.183)	(0.136)	(0.179)
Avg. Exp (Tch)	8.972	8.461	8.503	9.051	8.577
	(1.160)	(1.505)	(1.495)	(1.105)	(1.459)
N	107,550	227,444	1,936,790	200,156	2,556,132

Observations at student-year level. The mean of each characteristics is displayed with the standard deviations in parentheses. *% [Race] (Tch)* represents the average share of teachers for the given race.

Table 2: Summary of Student Noncognitive Measures by Race Match Status

	Asian		Black		Latino		White		Total	
	<50% RM	>50% RM	<50% RM	>50% RM	<50% RM	>50% RM	<50% RM	>50% RM	<50% RM	>50% RM
Share RM	0.122	0.565	0.146	0.665	0.200	0.644	0.306	0.714	0.186	0.659
	(0.130)	(0.105)	(0.145)	(0.155)	(0.144)	(0.145)	(0.122)	(0.157)	(0.147)	(0.151)
GPA	3.186	3.251	2.271	2.109	2.338	2.343	2.825	2.947	2.418	2.459
	(0.779)	(0.754)	(0.946)	(0.916)	(0.961)	(0.978)	(0.955)	(0.891)	(0.975)	(0.992)
Work Habits	1.617	1.648	1.081	0.962	1.151	1.178	1.408	1.467	1.192	1.227
	(0.427)	(0.413)	(0.523)	(0.491)	(0.531)	(0.533)	(0.530)	(0.504)	(0.539)	(0.542)
Cooperation	1.747	1.755	1.270	1.087	1.368	1.356	1.577	1.633	1.398	1.398
	(0.322)	(0.317)	(0.473)	(0.462)	(0.452)	(0.460)	(0.430)	(0.396)	(0.458)	(0.469)
Held Back	0.016	0.010	0.036	0.030	0.045	0.045	0.021	0.019	0.041	0.038
	(0.124)	(0.101)	(0.187)	(0.170)	(0.207)	(0.207)	(0.143)	(0.136)	(0.198)	(0.191)
Suspension	0.012	0.006	0.088	0.137	0.038	0.029	0.024	0.024	0.040	0.035
	(0.107)	(0.079)	(0.283)	(0.343)	(0.192)	(0.168)	(0.153)	(0.154)	(0.197)	(0.184)
Days Absent	4.268	3.617	9.837	11.630	8.385	7.859	8.070	7.887	8.164	8.085
	(7.758)	(7.959)	(12.639)	(14.326)	(12.630)	(11.613)	(10.569)	(9.893)	(12.262)	(11.505)
% Days Absent	0.025	0.022	0.061	0.076	0.051	0.048	0.048	0.046	0.050	0.049
	(0.048)	(0.051)	(0.085)	(0.099)	(0.083)	(0.076)	(0.067)	(0.063)	(0.080)	(0.075)
Learning Skills	0.799	0.852	-0.211	-0.457	-0.078	-0.056	0.411	0.529	-0.001	0.045
	(0.764)	(0.739)	(0.968)	(0.923)	(0.976)	(0.990)	(0.964)	(0.903)	(0.991)	(1.006)
Behavior	0.335	0.385	-0.130	-0.299	-0.001	0.038	0.104	0.122	0.019	0.036
	(0.588)	(0.575)	(1.042)	(1.162)	(1.026)	(0.970)	(0.787)	(0.739)	(0.995)	(0.944)
N	103,047	4,503	178,982	48,462	1,410,729	526,061	48,972	151,184	1,824,916	731,216

Observations at student-year level. The mean of each characteristics is displayed with the standard deviations in parentheses. *Share RM* is the share of race matched teachers a student has in a given year. *GPA* on 4-point scale. *Work Habits* and *Cooperation* are on a 2-point scale.

Table 3: Factor Loadings and Eigenvalues for Learning Skills and Behavior

Noncognitive Components	Learning Skills	Behavior
GPA	0.96	
Work Habits	0.98	
Cooperation	0.95	
Held Back		-0.67
% Days Absent		-0.76
Suspension		-0.45
Eigenvalue	2.78	1.23
Proportion	0.93	0.41

GPA: achievement grade point average; *Avg. Days Susp*: days suspended divided by times suspended; *Proportion* represents the share of the total variance explained by the components used to create each index.

Table 4: Student-Teacher Race Match Effects on Noncognitive Skills and Outcomes

Panel A Overall	(1) GPA	(2) Work Habits	(3) Cooperation	(4) Held Back	(5) Suspension	(6) ln(Days Absent)
Share RM	0.029***	0.037***	0.024***	0.003***	0.001	-0.025***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.005)
Panel B By Race	**(1) GPA**	**(2) Work Habits**	**(3) Cooperation**	**(4) Held Back**	**(5) Suspension**	**(6) ln(Days Absent)**
Asian × Share RM	0.080***	0.081***	0.071***	-0.010***	-0.006*	-0.099***
	(0.009)	(0.009)	(0.009)	(0.002)	(0.003)	(0.017)
Black × Share RM	0.074***	0.071***	0.071***	0.001	0.003	-0.067***
	(0.007)	(0.007)	(0.007)	(0.002)	(0.004)	(0.011)
Latino × Share RM	0.022***	0.036***	0.021***	0.006***	0.001	-0.001
	(0.005)	(0.005)	(0.005)	(0.001)	(0.002)	(0.007)
White × Share RM	0.002	0.011	-0.019***	0.005**	0.001	-0.008
	(0.006)	(0.006)	(0.005)	(0.002)	(0.002)	(0.009)
N	2,499,655	2,499,655	2,499,655	2,499,655	2,499,655	2,499,655

Student-clustered standard errors shown in parentheses. All columns include school-grade, year, and student fixed effects. All models include time-variant student, teacher, and class-averaged characteristics. *ln(Days Absent)* is calculated using ln(x+1) to account for zeros.

Table 5: Student-Teacher Race Match Effects on Learning Skills and Behavior Indices

Panel A Overall	(1) Learning Skills Index	(2) Behavior Index
Share RM	0.011**	0.041***
	(0.004)	(0.003)
Panel B By Race	**(1) Learning Skills Index**	**(2) Behavior Index**
Asian × Share RM	0.137***	0.058***
	(0.012)	(0.011)
Black × Share RM	0.096***	0.046***
	(0.009)	(0.013)
Latino × Share RM	0.038***	0.001
	(0.006)	(0.007)
White × Share RM	-0.014	-0.009
	(0.007)	(0.008)
N	2,499,655	2,499,655

Student-clustered standard errors shown in parentheses. All columns include school-grade, year, and student fixed effects. All models include time-variant student, teacher, and class-averaged characteristics.

Table 6: Race Match Effects on Reference Bias Adjusted Learning Skills and Behavior Indices

	(1) LSI	(2) C1: LSI	(3) C2: LSI	(4) BI	(5) C1: BI	(6) C2: BI
Asian × Share RM	0.119***	0.127***	0.126***	0.066***	0.028	0.072***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)
Black × Share RM	0.105***	0.095***	0.088***	0.041*	0.068***	0.060***
	(0.011)	(0.011)	(0.011)	(0.017)	(0.018)	(0.018)
Latino × Share RM	0.039***	0.009	0.010	-0.013	-0.016	0.010
	(0.007)	(0.007)	(0.007)	(0.009)	(0.010)	(0.010)
White × Share RM	-0.013	-0.012	-0.012	0.004	-0.016	-0.023*
	(0.009)	(0.009)	(0.009)	(0.010)	(0.011)	(0.010)
N	1,593,764	1,593,764	1,593,764	1,593,764	1,593,764	1,593,764

Student-clustered standard errors shown in parentheses. All columns include school-grade, year, and student fixed effects. All models include time-variant student, teacher, and class-averaged characteristics.

Table 7: Factor Loadings and Sampling Adequacy for Exploratory Factor Analysis

	Learning Skills (EFA)	Cognitive	Behavior (EFA)	Uniqueness
GPA	0.80	0.38	0.33	0.10
Work Habits	0.87	0.31	0.29	0.06
Cooperation	0.79	0.34	0.28	0.19
Held Back	-0.04	-0.05	-0.26	0.93
% Days Absent	-0.24	-0.12	-0.38	0.79
Suspension	-0.15	-0.05	-0.20	0.94
Test ELA	0.28	0.72	0.16	0.38
Test Math	0.25	0.72	0.18	0.39
Eigenvalue	2.24	1.42	0.58	

EFA: exploratory factor analysis. This table represents a single EFA where each factor (represented by the column headings) was determined using the factor loadings of each noncognitive component.

Table 8: Race Match Effects on Confirmatory and Exploratory Factor Analysis Generated LSI

	(1) Learning Skills (CFA)	(2) Learning Skills (EFA)
Share RM	0.011***	0.066***
	(0.004)	(0.007)

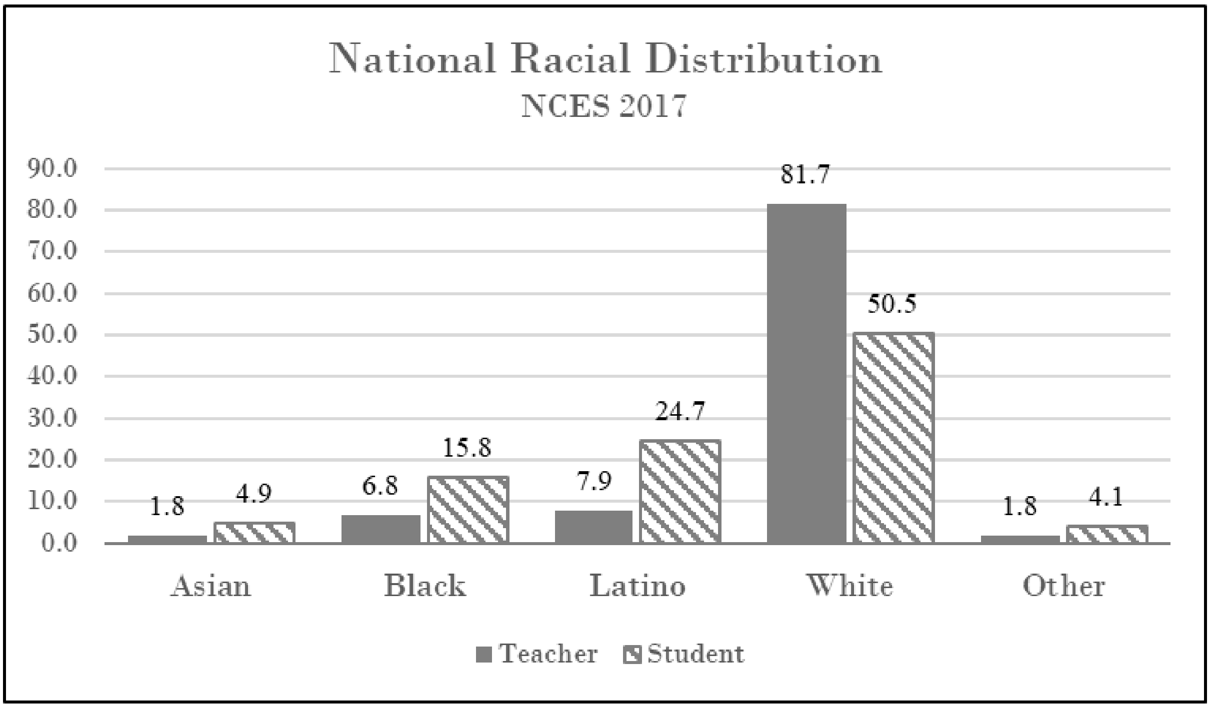
	(1) Learning Skills (CFA)	(2) Learning Skills (EFA)
Asian × Share RM	0.137***	0.079***
	(0.012)	(0.024)
Black × Share RM	0.096***	0.126***
	(0.009)	(0.020)
Latino × Share RM	0.038***	0.089***
	(0.006)	(0.013)
White × Share RM	-0.014	-0.004
	(0.007)	(0.015)
N	2,499,655	1,572,850

Student-clustered standard errors shown in parentheses. All columns include school-grade, year, and student fixed effects. All models include time-variant student, teacher, and class-averaged characteristics. CFA: confirmatory factor analysis; EFA: exploratory factor analysis.

Figures

Figure 1: Teacher and Student Racial Distributions

(a) National Distributions



(b) LAUSD Distributions (from analytical sample)

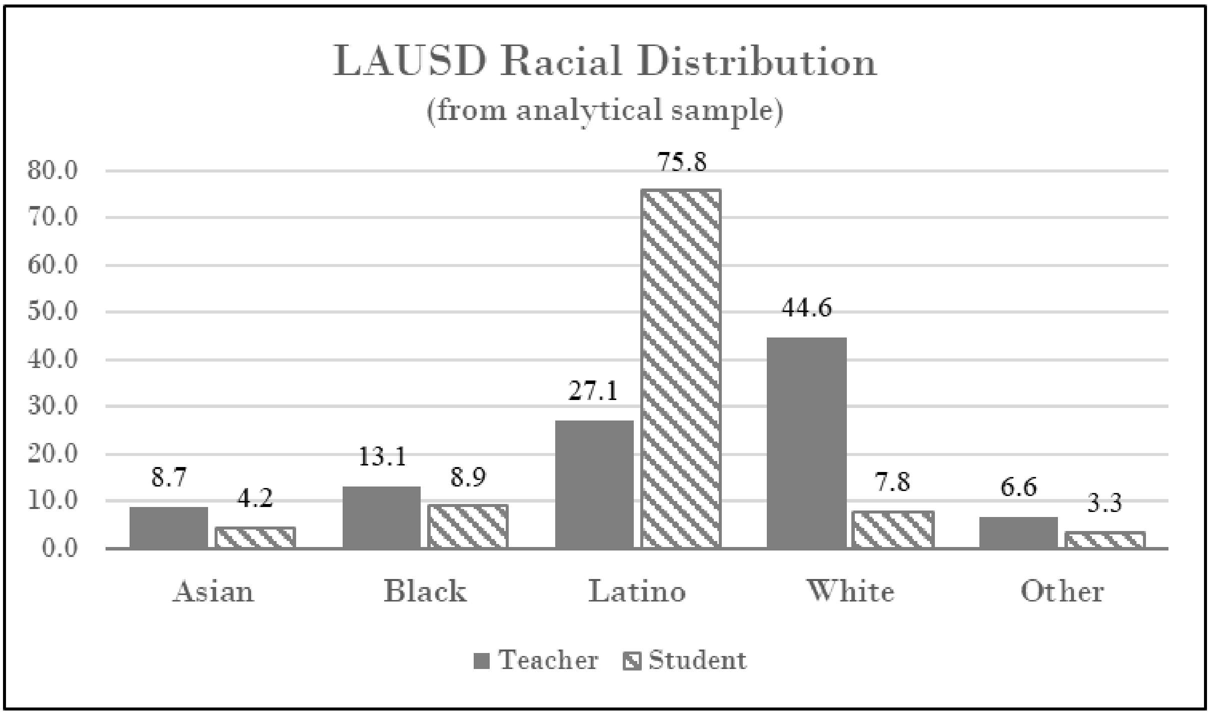


Figure 2: Average Percent (%) Race Match by Student Race

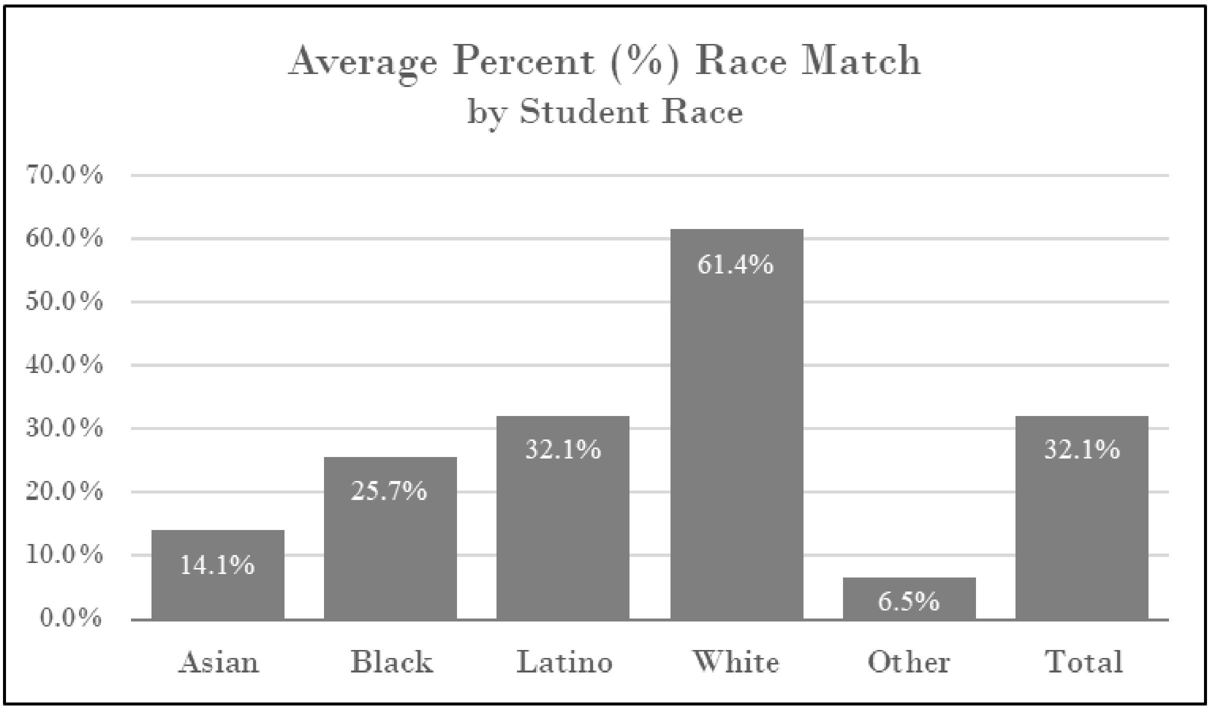
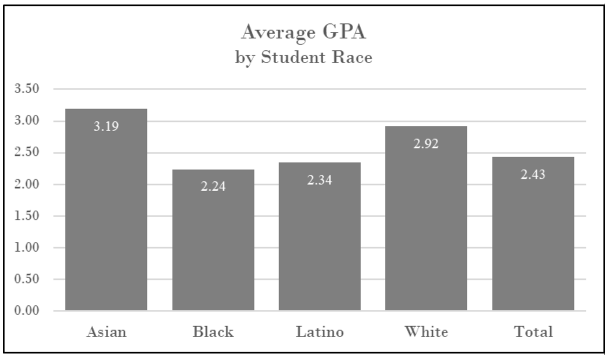
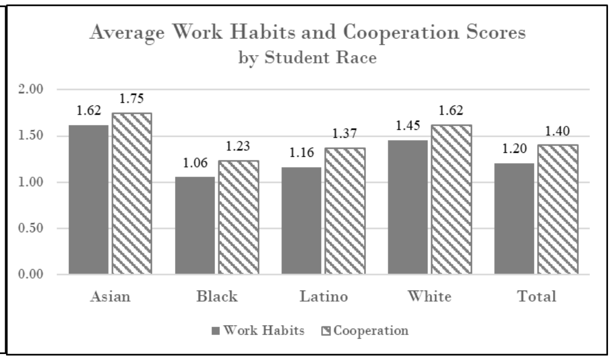


Figure 3: Summary of Noncognitive Outcomes per Year by Race

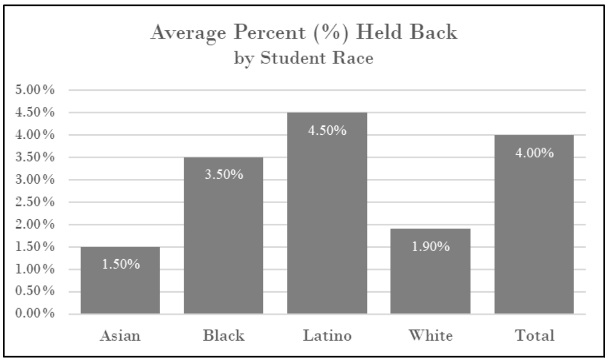
(a) Average GPA



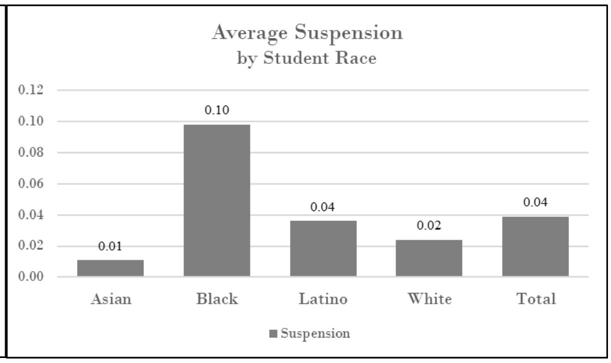
(b) Average Work Habits and Cooperation Scores



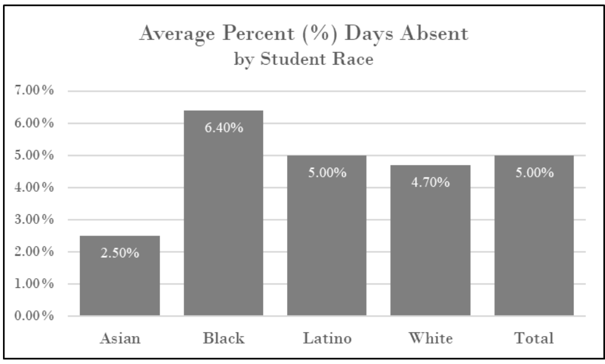
(c) Average Percent (%) Held Back



(d) Average Suspensions



(e) Average Percent (%) Days Absent



(f) Standardized Learning Skills and Behavior

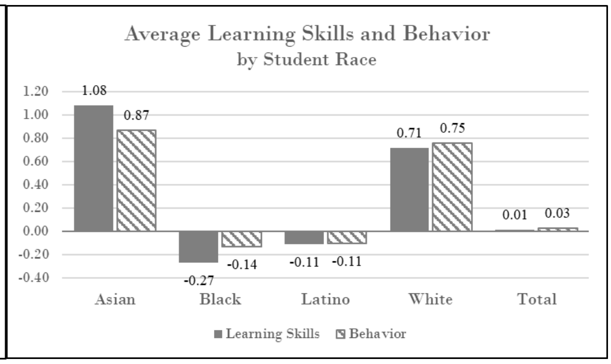


Figure 4: Race Match Effect on Standardized Noncognitive Measures

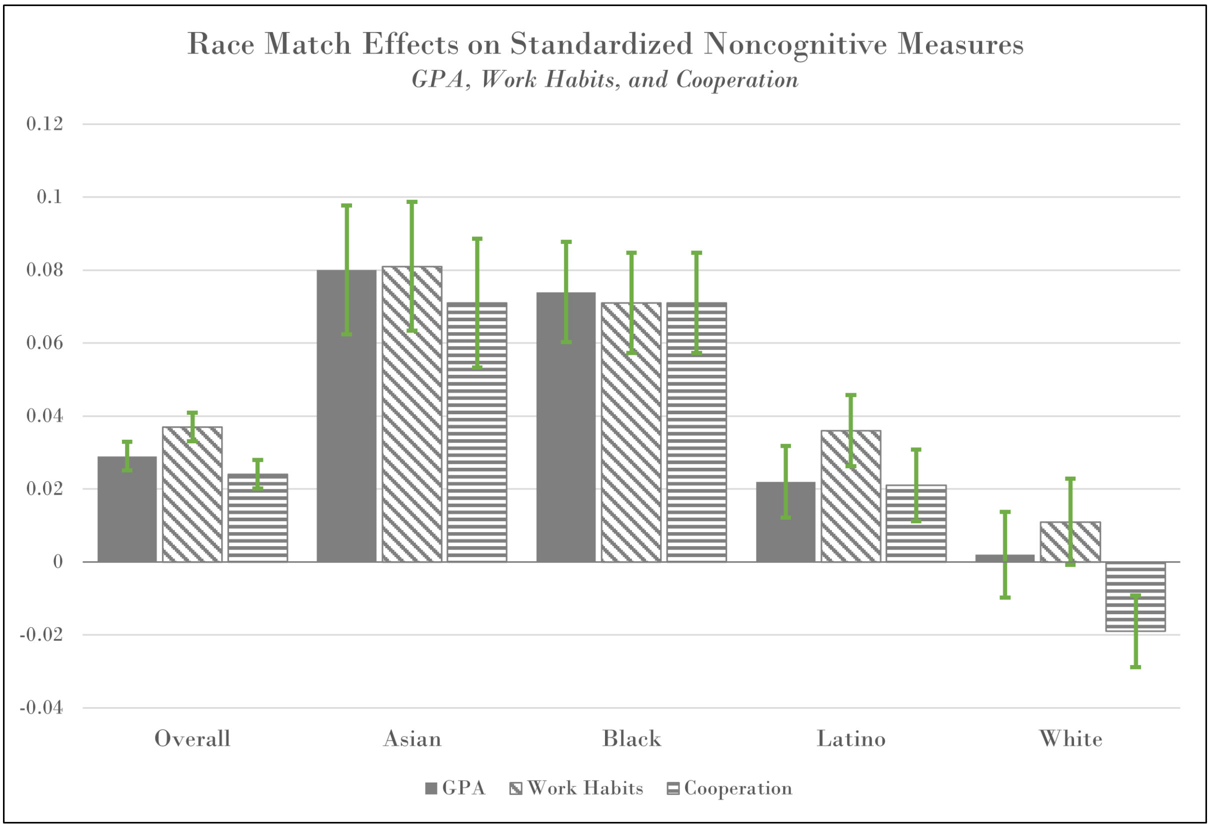


Figure 5: Race Match Effect on the Probability of Noncognitive Measures

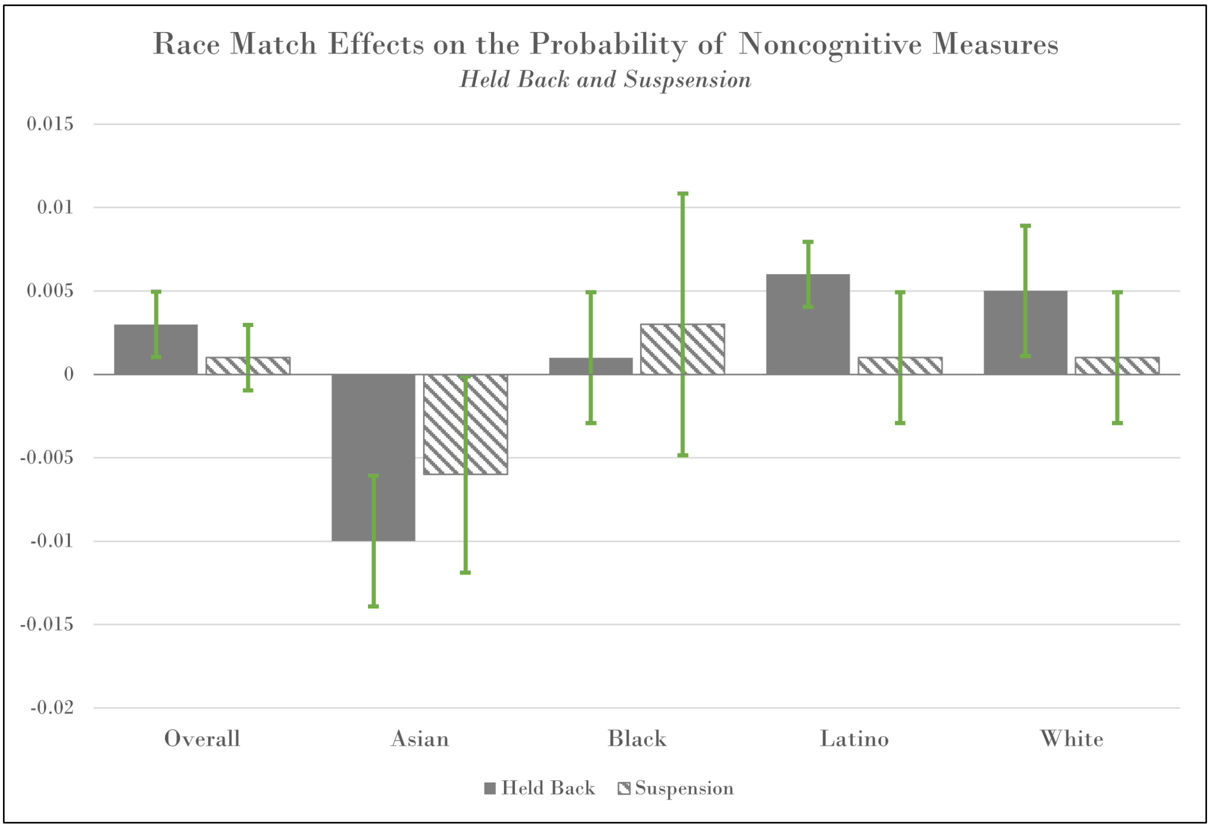


Figure 6: Race Match Effect on ln(Days Absent)

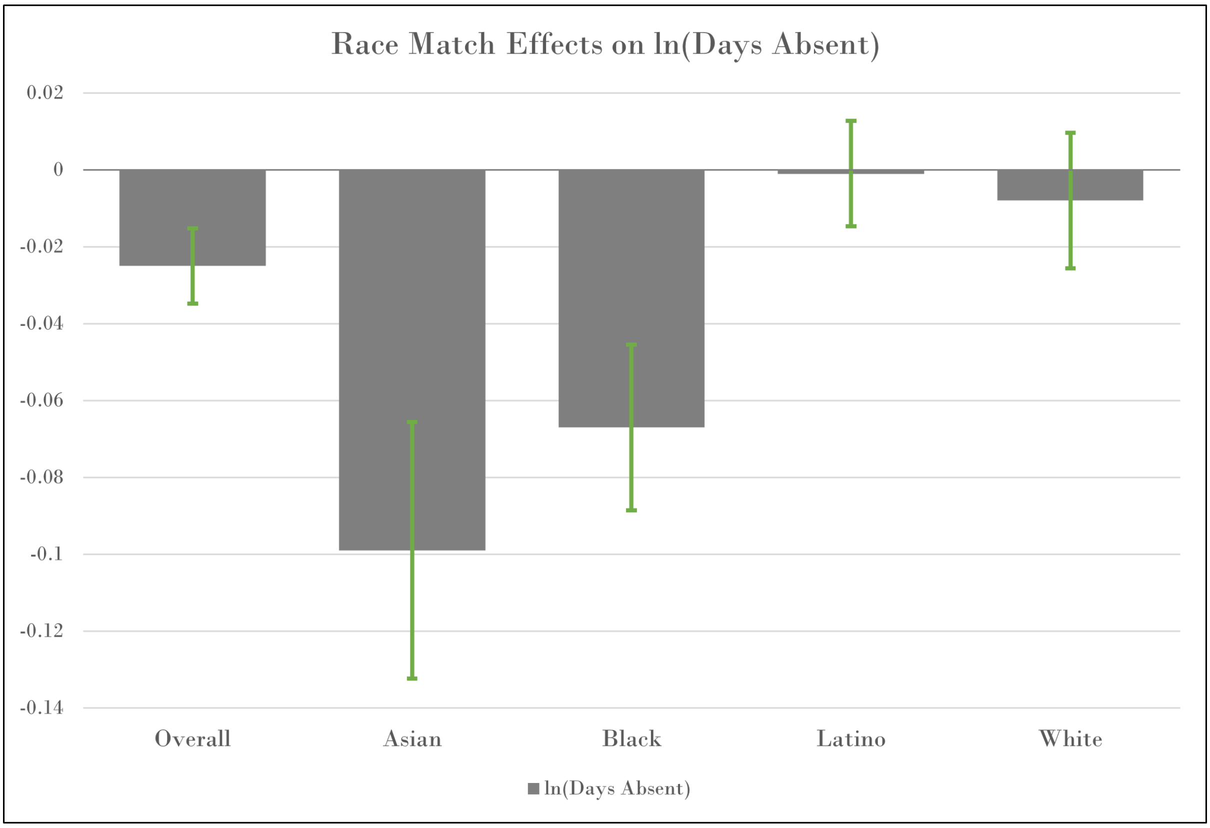


Figure 7: Race Match Effect on Learning Skills and Behavior

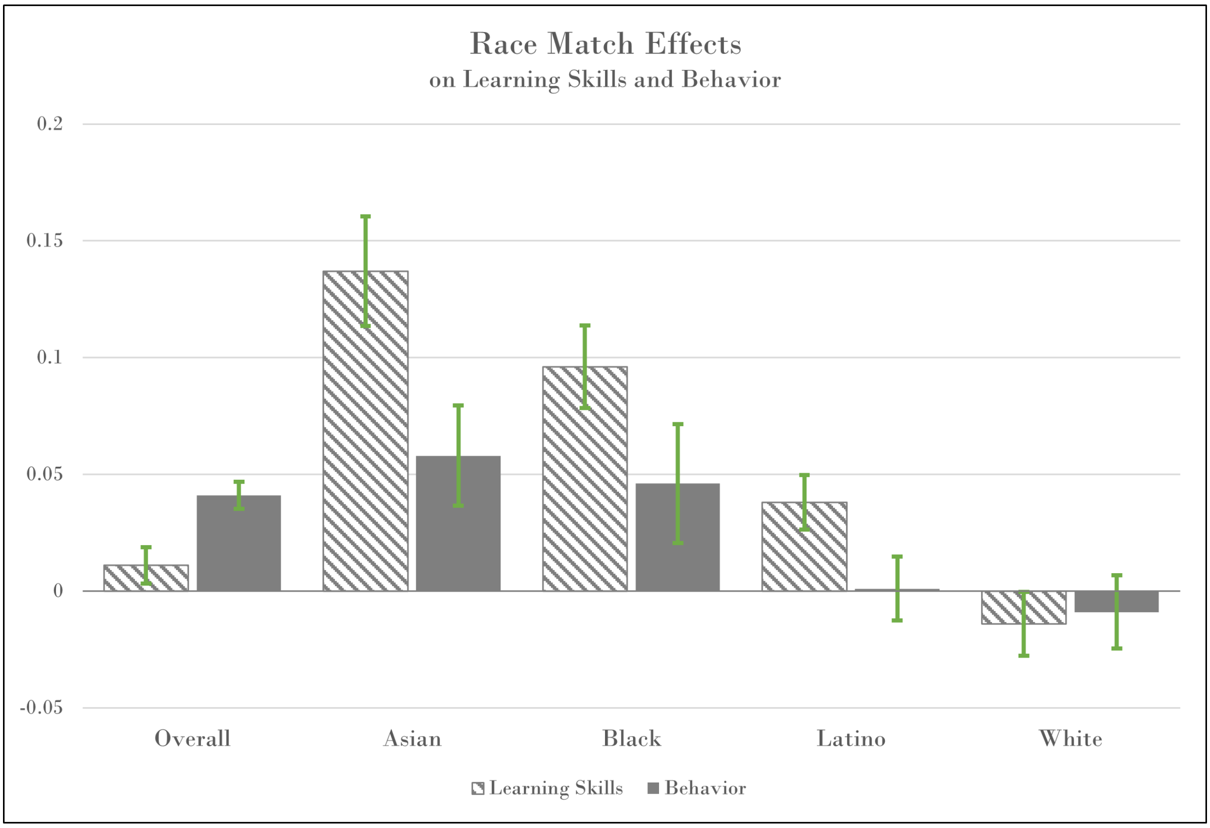
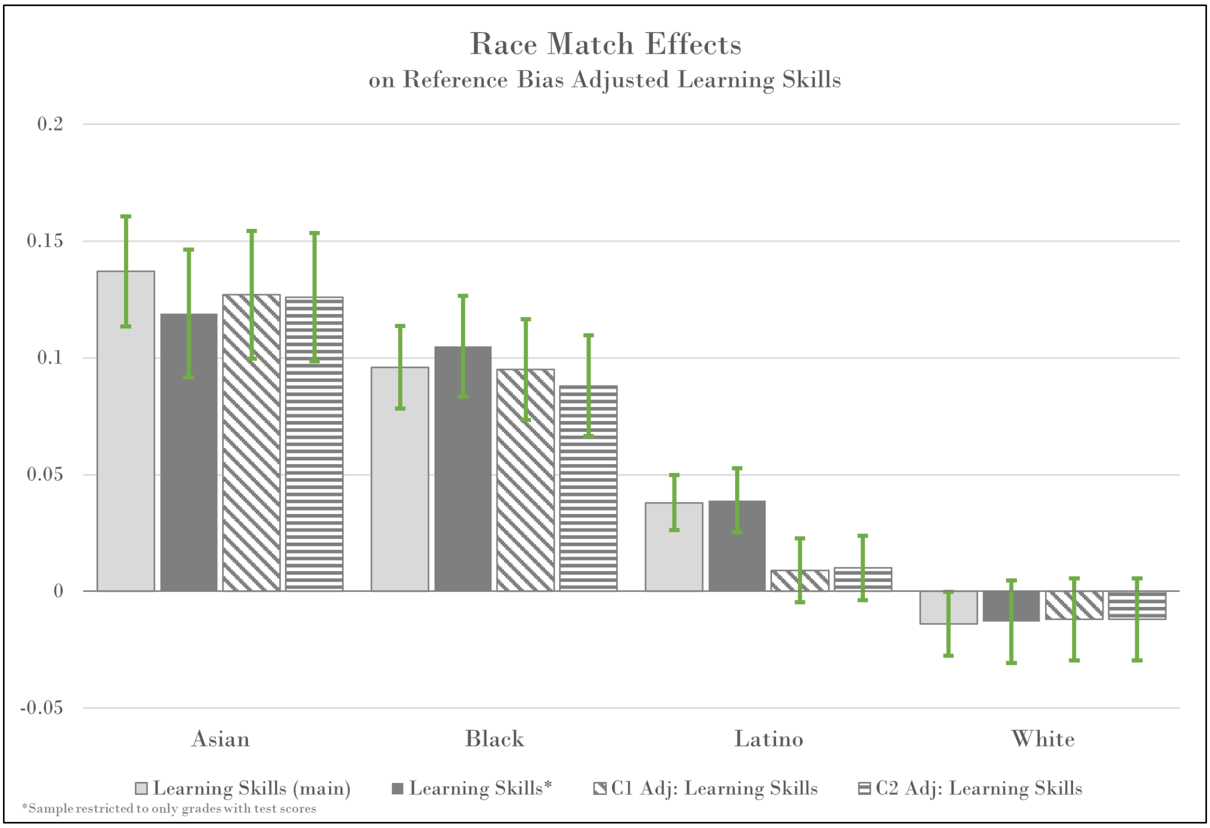


Figure 8: Race Match Effect on Reference Bias Adjusted Learning Skills and Behavior

(a) Learning Skills



(b) Behavior

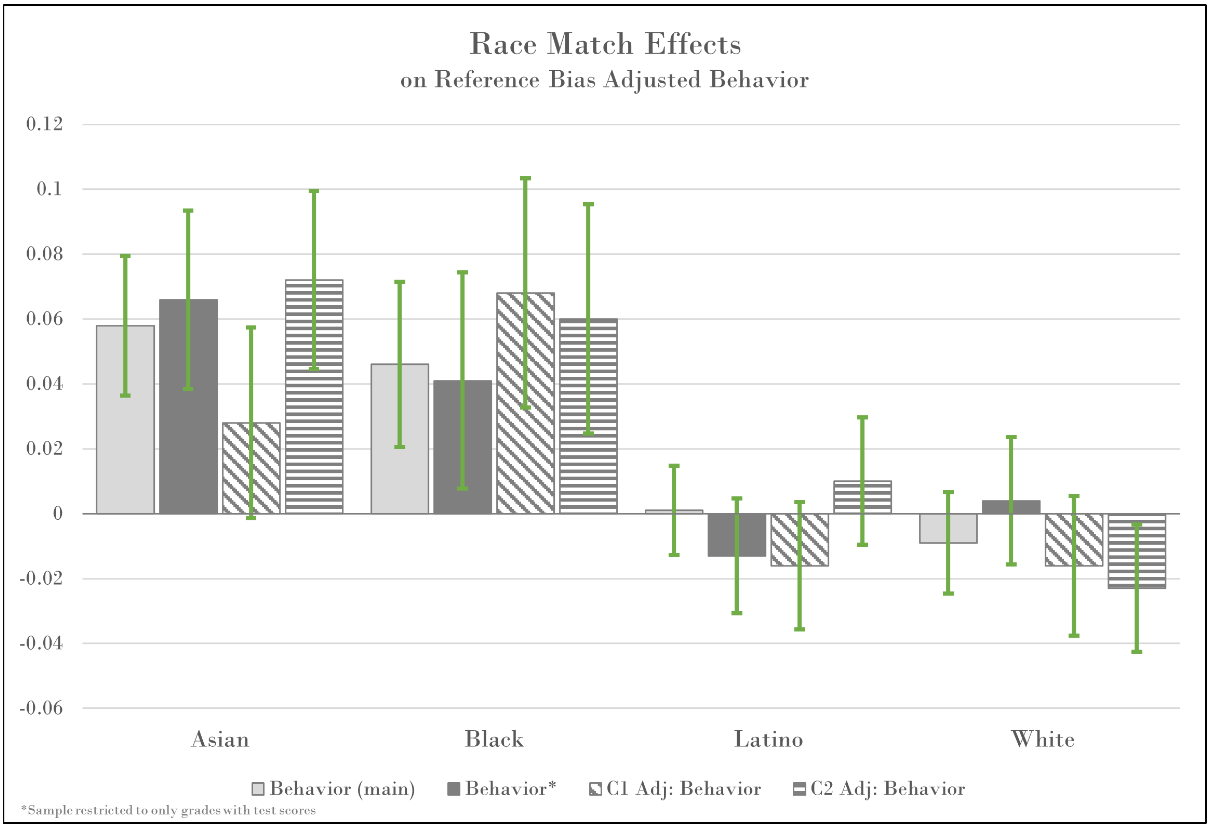


Figure 9: Course-Level Race Match Effect on GPA

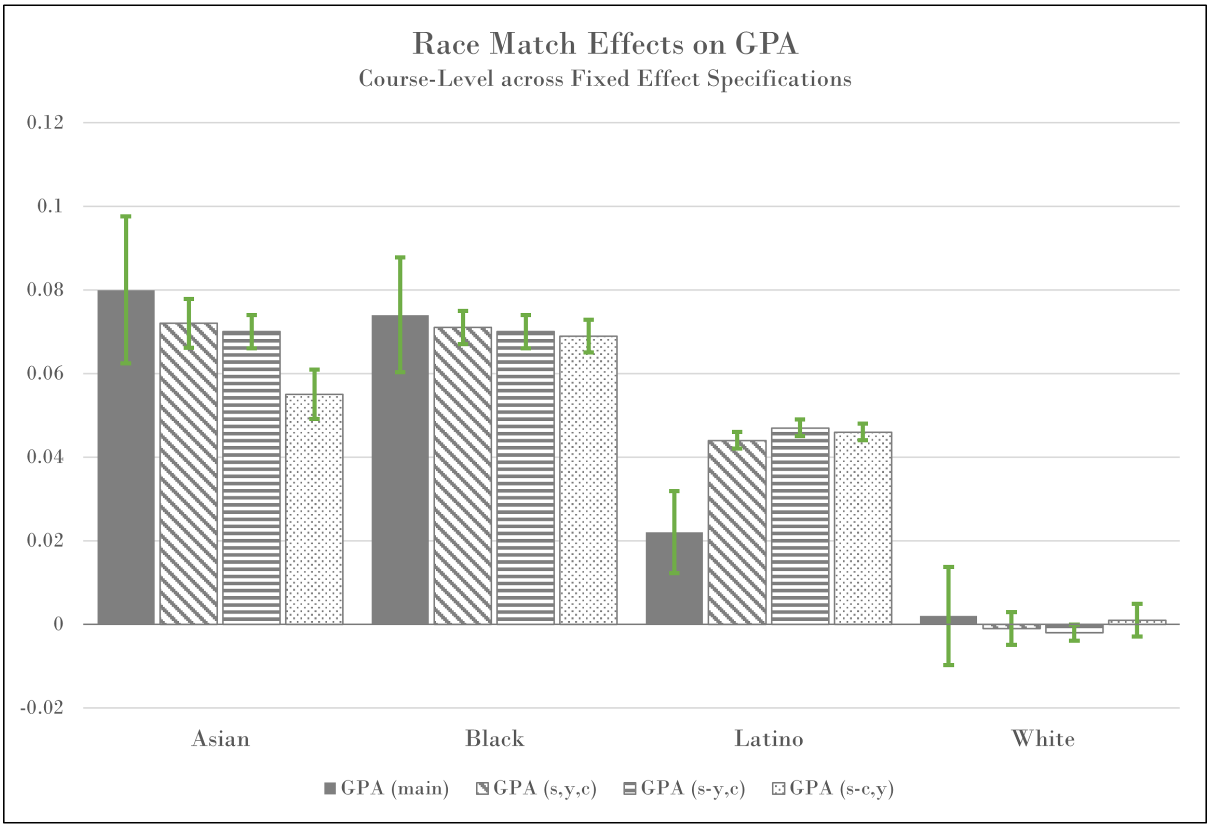


Figure 10: Course-Level Race Match Effect on Work Habits

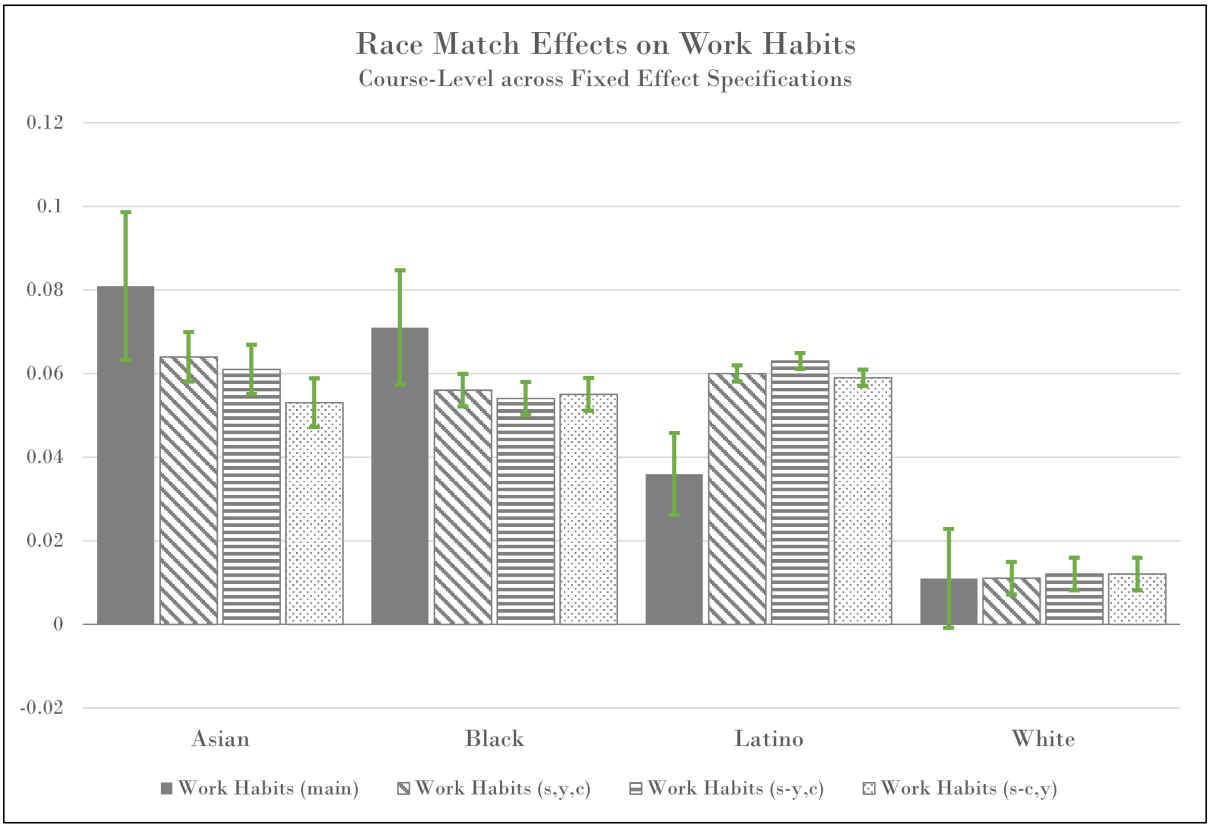


Figure 11: Course-Level Race Match Effect on Cooperation

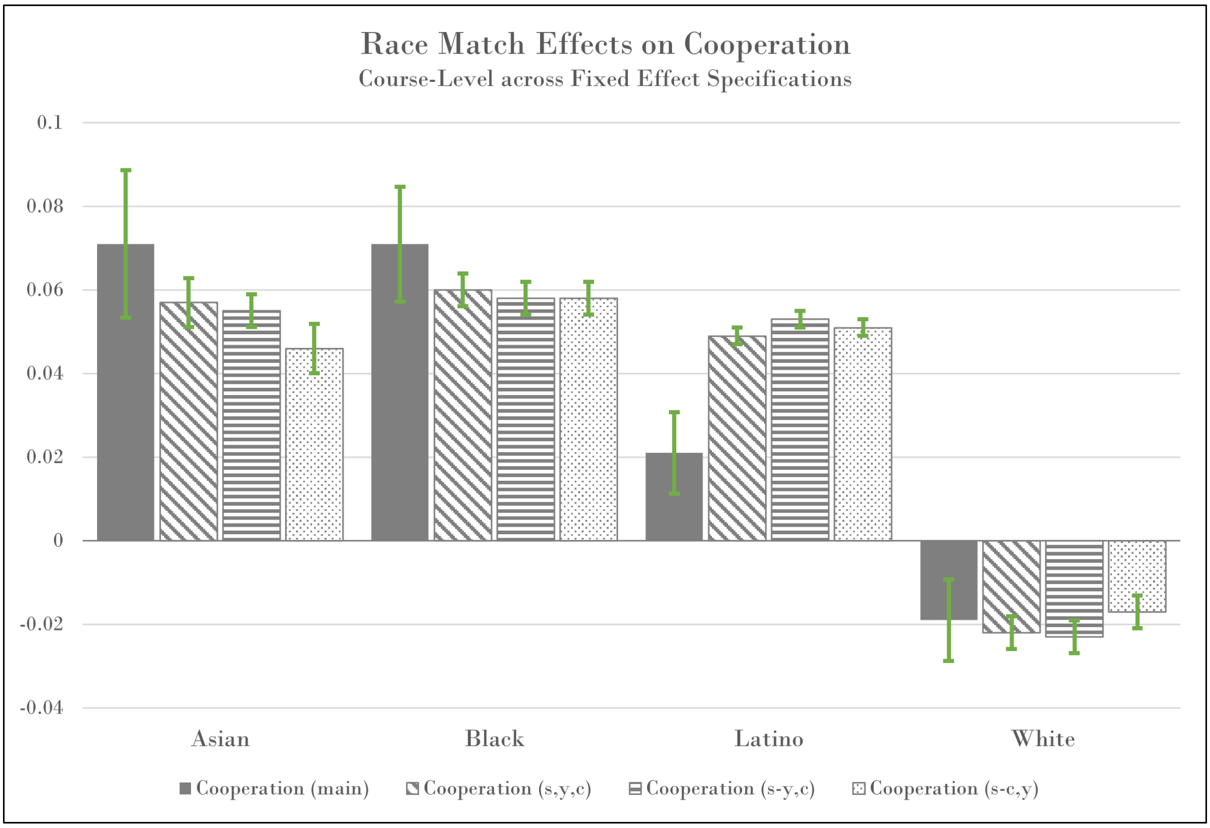


Figure 12: Course-Level Race Match Effect on Learning Skills Index

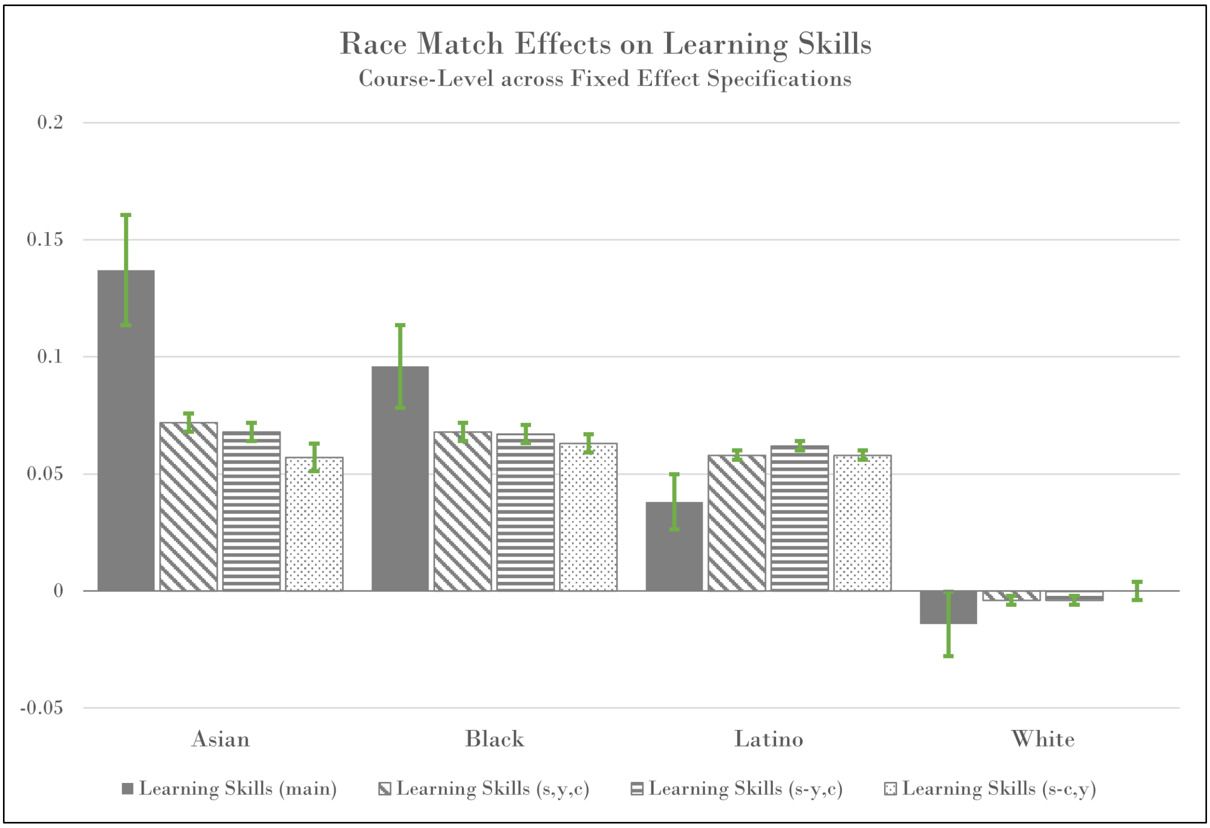


Figure 13: Race Match Effect with Lagged Outcome in Model

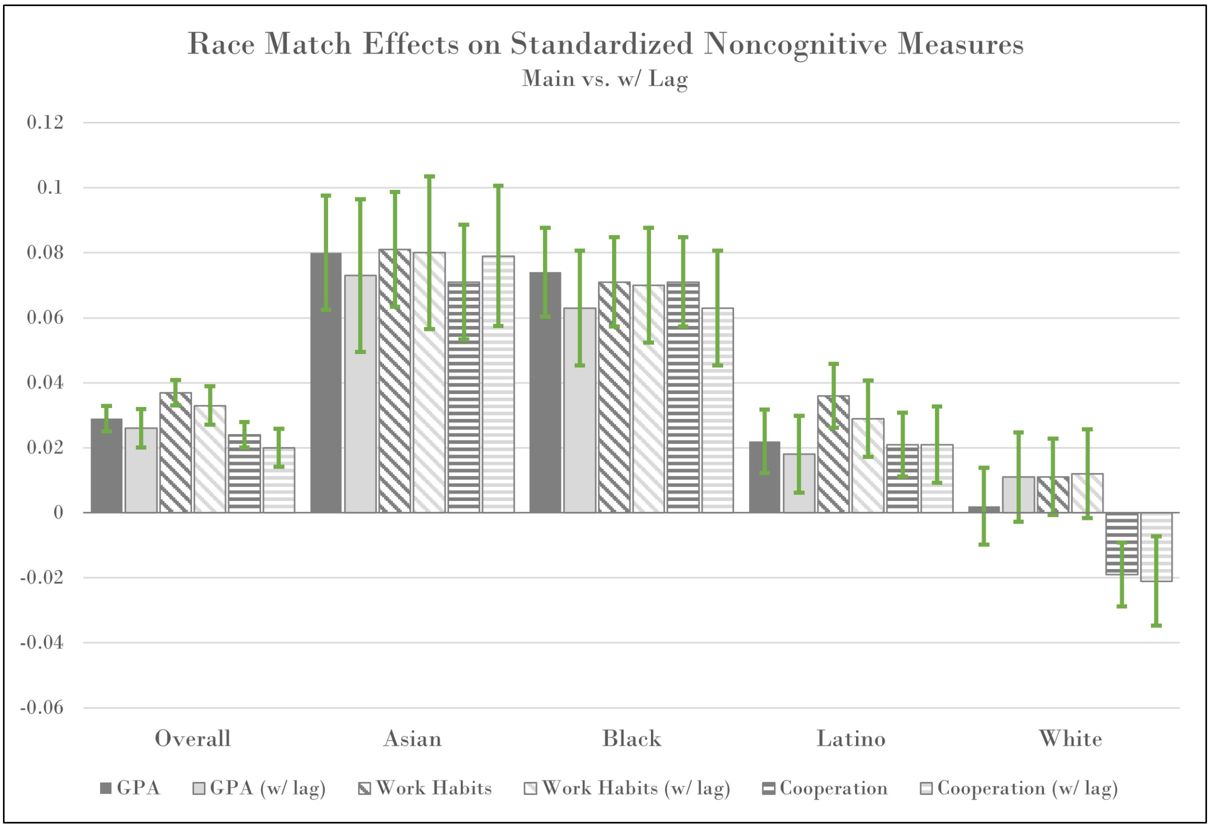


Figure 14: Race Match Effect with Lagged Outcome in Model

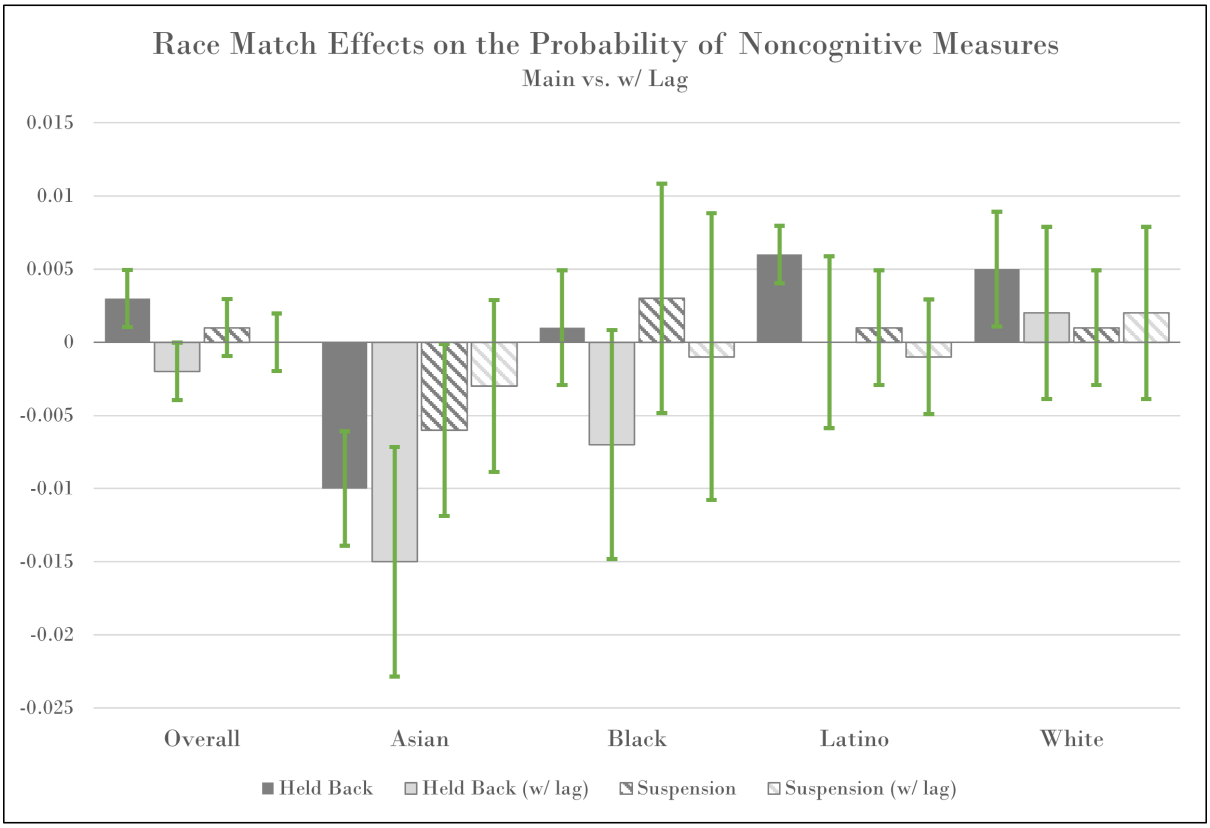


Figure 15: Race Match Effect with Lagged Outcome in Model

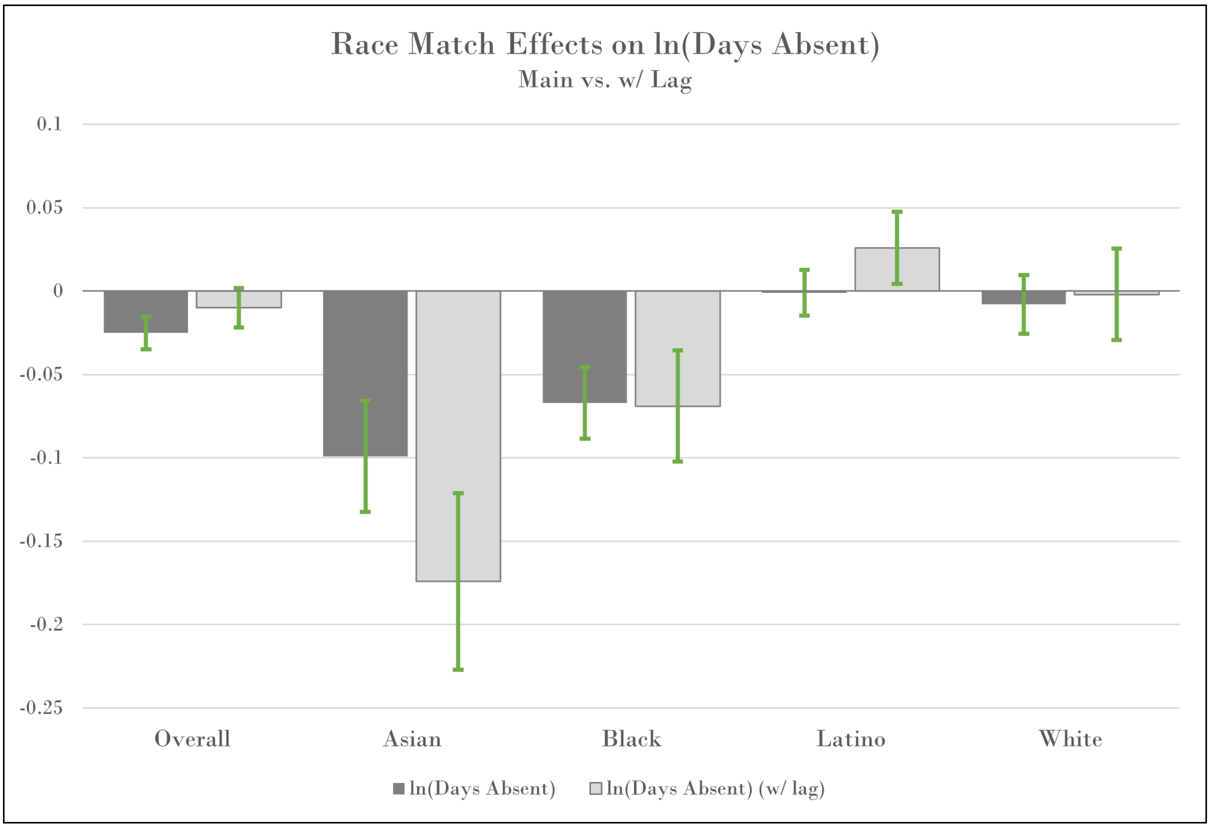


Figure 16: Race Match Effect with Lagged Outcome in Model

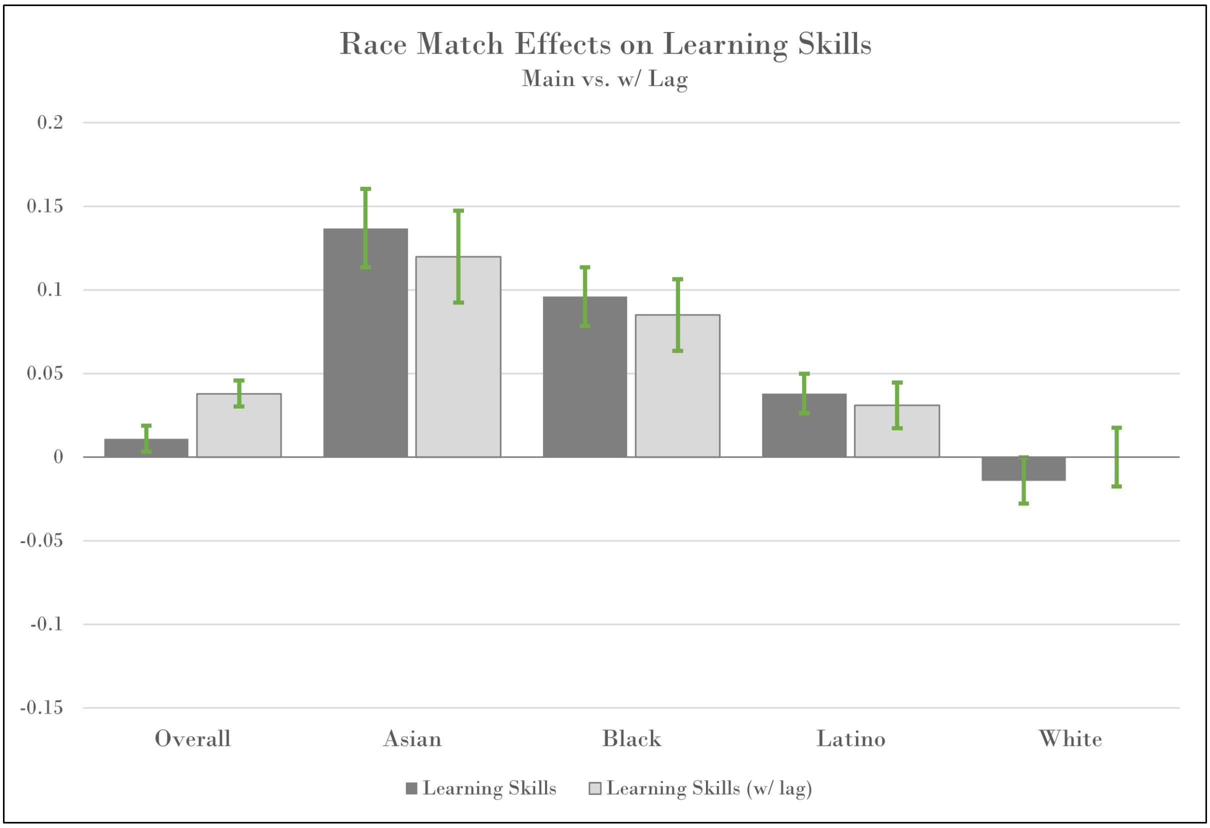


Figure 17: Race Match Effect with Lagged Outcome in Model

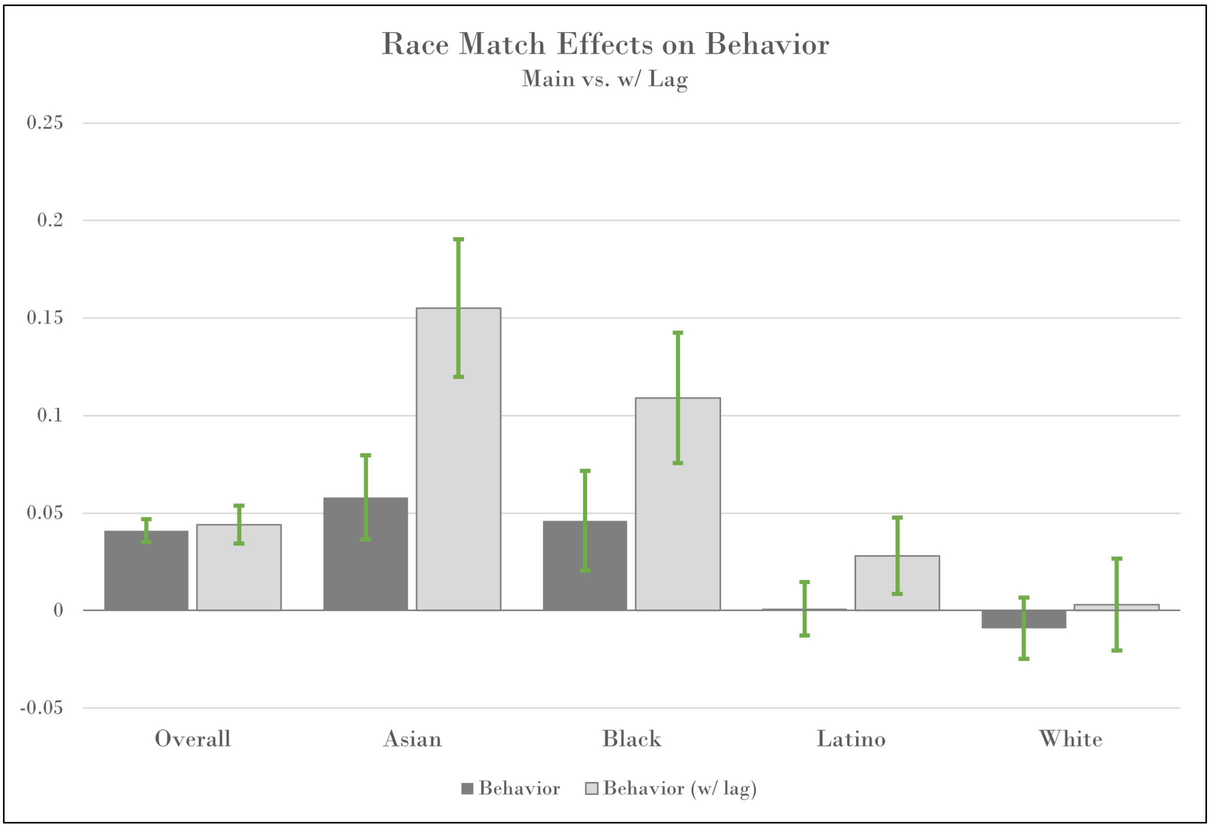
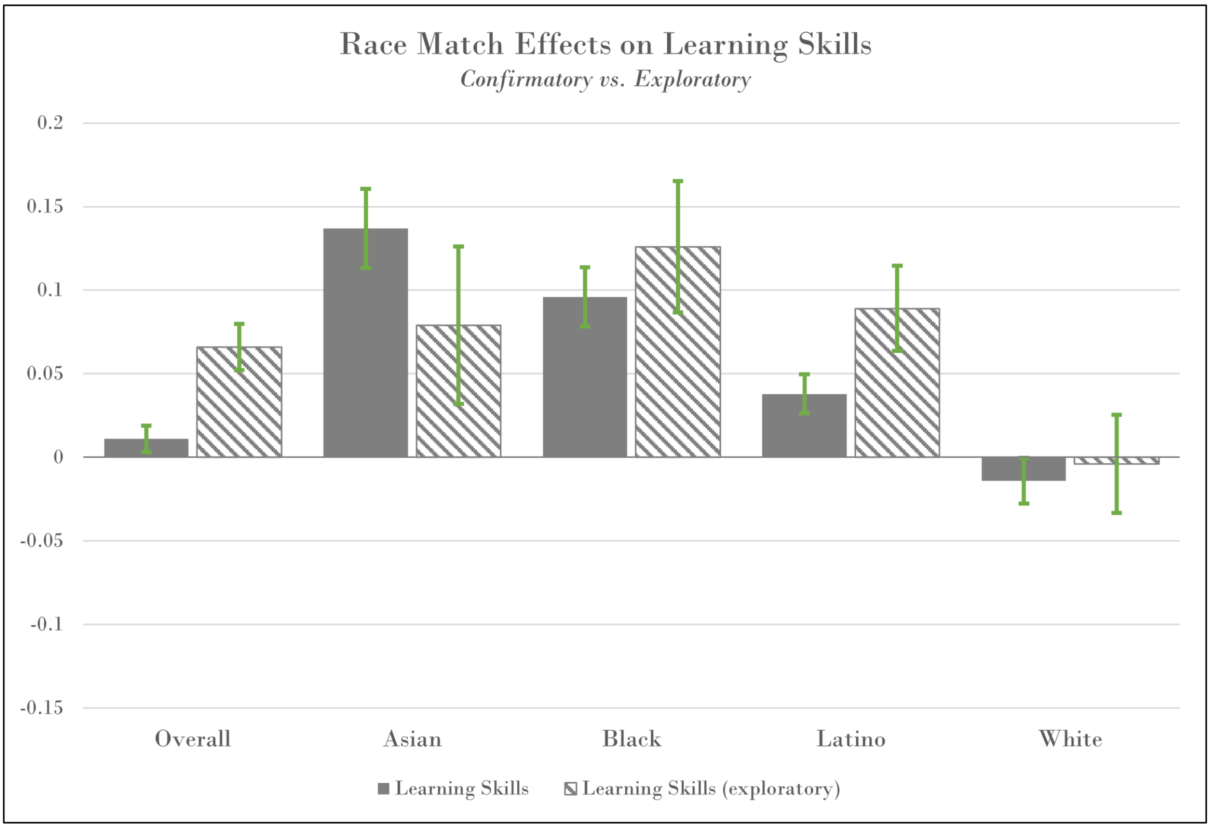


Figure 18: Race Match Effect using Confirmatory vs Exploratory Factor Analysis



Appendix A: Student Descriptive Characteristics and Noncognitive Skill Correlations

Table A1: Summary of Student Noncognitive Skills and Outcomes

	Asian	Black	Latino	White	Total
Share RM	0.141	0.257	0.320	0.614	0.322
	(0.157)	(0.258)	(0.245)	(0.230)	(0.260)
GPA	3.189	2.237	2.339	2.917	2.430
	(0.778)	(0.942)	(0.966)	(0.908)	(0.980)
Work Habits	1.618	1.055	1.158	1.452	1.202
	(0.427)	(0.518)	(0.531)	(0.511)	(0.540)
Cooperation	1.747	1.231	1.365	1.619	1.398
	(0.322)	(0.477)	(0.454)	(0.405)	(0.461)
Held Back	0.015	0.035	0.045	0.019	0.040
	(0.123)	(0.184)	(0.207)	(0.138)	(0.196)
Suspension	0.011	0.098	0.036	0.024	0.039
	(0.106)	(0.297)	(0.186)	(0.154)	(0.193)
Days Absent	4.241	10.219	8.242	7.932	8.142
	(7.767)	(13.037)	(12.364)	(10.063)	(12.051)
%Days Absent	0.025	0.064	0.050	0.047	0.050
	(0.048)	(0.088)	(0.081)	(0.064)	(0.079)
Learning Skills	0.801	-0.263	-0.072	0.500	0.012
	(0.763)	(0.964)	(0.980)	(0.919)	(0.996)
Behavior	0.337	-0.166	0.009	0.118	0.023
	(0.588)	(1.071)	(1.011)	(0.751)	(0.981)
N	107,550	227,444	1,936,790	200,156	2,556,132

Observations at student-year level. The mean of each characteristics is displayed with the standard deviations in parentheses. *Share RM* is the share of race matched teachers a student has in a given year. *GPA* on 4-point scale. *Work Habits* and *Cooperation* are on a 2-point scale.

Table A2: Cross-correlations of Noncognitive Components

Variables	GPA	Work Habits	Cooperation	Held Back	Absent	Suspension
GPA	1.000					
Work Habits	0.936	1.000				
Cooperation	0.844	0.895	1.000			
Held Back	-0.198	-0.181	-0.167	1.000		
Absent	-0.369	-0.343	-0.333	0.184	1.000	
Suspension	-0.173	-0.180	-0.209	0.027	0.108	1.000

Appendix B: Reference Bias

Using the methods found in Elder and Zhou (2021), I adjust each learning skills measure for teacher reference bias in order to account for the subjective nature of noncognitive skills. Specifically, I generate the normalized skill level of a specific school, grade, year, and student race, C_{sgyr}, for math test scores using the following equation:

$$C_{sgyr} = (\hat{\mu}^{math}_{sgyr} - \hat{\mu}^{math}_{gyr})/\hat{\sigma}^{math}_{gyr}. \tag{9}$$

Under condition 1, for each learning skill measure, $k \in \{\text{GPA, Effort, Coop}\}$, I calculate the expected mean of k when adjusted by the normalized skill level using:

$$\hat{\mu}^{k}_{sgyr} = C_{sgyr} * \hat{\sigma}^{k}_{gyr} + \hat{\mu}^{k}_{gyr}. \tag{10}$$

I calculate the reference bias adjusted learning skill measure for student i by subtracting the reference bias from the original learning skill measure:

$$\hat{Y}^{Adjust,k}_{igyr} = Y^{Orig,k}_{igyr} - (\mu^{k}_{sgyr} - \hat{\mu}^{k}_{sgyr}). \tag{11}$$

The final adjustment is to scale the newly calculated outcomes so that the variance matches that of the original outcomes:

$$Y^{Adjust,k}_{igyr} = \hat{Y}^{Adjust,k}_{igyr} * (\hat{\sigma}^{k}_{gyr}/\hat{\sigma}^{k}_{sgyr}). \tag{12}$$

For condition 2, I repeat the steps above and use GPA as the given skill k in equation 10. Then, the reference bias adjusted outcome for a given learning skill l under condition 2 is as follows:

$$\hat{Y}^{Adjust,l}_{igyr} = Y^{Orig,l}_{igyr} - (\mu^{GPA}_{sgyr} - \hat{\mu}^{GPA}_{sgyr}). \tag{13}$$

Again, I adjust $\hat{Y}^{Adjust,l}_{igyr}$ so that the variance matches that of $\hat{Y}^{Orig,l}_{igyr}$:

$$Y^{Adjust,l}_{igyr} = \hat{Y}^{Adjust,l}_{igyr} * (\hat{\sigma}^{l}_{gyr} / \hat{\sigma}^{GPA}_{sgyr}). \tag{14}$$

Summary of Noncognitive Skills for Math Test Restricted Sample

Table B1 summarizes the original and reference bias adjusted noncognitive skill components for the sample restricted to students that have a math test score.

Table B1: Reference Bias Adjusted Outcomes

	Asian		Black		Latino		White	
	<50% RM	>50% RM	<50% RM	>50% RM	<50% RM	>50% RM	<50% RM	>50% RM
GPA	3.222	3.274	2.272	2.097	2.339	2.337	2.856	2.965
	(0.761)	(0.723)	(0.928)	(0.894)	(0.949)	(0.962)	(0.939)	(0.881)
C1 Bias: GPA	3.089	3.139	2.226	1.871	2.292	2.254	2.799	2.927
	(0.762)	(0.708)	(0.926)	(0.857)	(0.950)	(0.957)	(0.932)	(0.883)
C2 Bias: GPA	3.222	3.275	2.313	1.944	2.349	2.310	2.840	2.971
	(0.795)	(0.738)	(0.962)	(0.891)	(0.974)	(0.981)	(0.946)	(0.896)
Work Habits	1.629	1.655	1.074	0.951	1.146	1.168	1.420	1.471
	(0.419)	(0.401)	(0.517)	(0.484)	(0.526)	(0.528)	(0.524)	(0.502)
C1 Bias: Work Habits	1.571	1.595	1.051	0.855	1.127	1.120	1.384	1.449
	(0.420)	(0.389)	(0.517)	(0.468)	(0.527)	(0.525)	(0.521)	(0.503)
C2 Bias: Work Habits	1.629	1.656	1.097	0.866	1.151	1.153	1.410	1.474
	(0.441)	(0.413)	(0.543)	(0.489)	(0.545)	(0.542)	(0.533)	(0.513)
Cooperation	1.753	1.760	1.260	1.067	1.361	1.345	1.585	1.635
	(0.314)	(0.306)	(0.471)	(0.457)	(0.447)	(0.454)	(0.423)	(0.393)
C1 Bias: Cooperation	1.710	1.723	1.250	1.055	1.345	1.327	1.564	1.622
	(0.315)	(0.294)	(0.474)	(0.448)	(0.448)	(0.454)	(0.417)	(0.395)
C2 Bias: Cooperation	1.753	1.760	1.281	0.990	1.366	1.332	1.578	1.637
	(0.334)	(0.314)	(0.508)	(0.465)	(0.474)	(0.470)	(0.433)	(0.404)
Held Back	0.011	0.005	0.027	0.021	0.034	0.034	0.015	0.011
	(0.104)	(0.072)	(0.162)	(0.143)	(0.182)	(0.181)	(0.122)	(0.106)
C1 Bias: Held Back	0.010	0.007	0.032	-0.003	0.035	0.028	0.009	0.013
	(0.104)	(0.077)	(0.163)	(0.136)	(0.182)	(0.181)	(0.121)	(0.107)
C2 Bias: Held Back	0.011	0.006	0.033	-0.000	0.035	0.031	0.014	0.012
	(0.111)	(0.081)	(0.175)	(0.148)	(0.192)	(0.187)	(0.128)	(0.111)
Suspension	0.014	0.006	0.104	0.162	0.046	0.036	0.028	0.027
	(0.116)	(0.080)	(0.305)	(0.368)	(0.210)	(0.186)	(0.166)	(0.163)
C1 Bias: Suspension	0.013	0.008	0.131	0.032	0.046	0.030	0.017	0.029
	(0.116)	(0.085)	(0.311)	(0.340)	(0.210)	(0.184)	(0.165)	(0.163))
C2 Bias: Suspension	0.014	0.006	0.117	0.111	0.048	0.030	0.025	0.028
	(0.124)	(0.090)	(0.320)	(0.378)	(0.220)	(0.193)	(0.175)	(0.170))
N	72,087	3,252	117,286	31,672	951,934	348,296	34,125	106,185

Observations at student-year level. The mean of each characteristics is displayed with the standard deviations in parentheses. *C1 Bias* refers to condition 1 as the reference bias adjustment method. *C2 Bias* refers to condition 2 as the reference bias adjustment method.

Appendix C: Robustness Checks on Sorting

Course Analysis

Table C1: Fixed Effect Level Course Analysis on Learning Skills Components

	(1) GPA	(2) GPA	(3) GPA	(4) Work Habits	(5) Work Habits	(6) Work Habits	(7) Cooperation	(8) Cooperation	(9) Cooperation
RM × Asian	0.072***	0.070***	0.055***	0.064***	0.061***	0.053***	0.057***	0.055***	0.046***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
RM × Black	0.071***	0.070***	0.069***	0.056***	0.054***	0.055***	0.060***	0.058***	0.058***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
RM × Latino	0.044***	0.047***	0.046***	0.060***	0.063***	0.059***	0.049***	0.053***	0.051***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
RM × White	-0.001	-0.002	0.001	0.011***	0.012***	0.012***	-0.022***	-0.023***	-0.017***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N	16337360	16318494	15805680	16337360	16318494	15805680	16337360	16318494	15805680
Student FE	X			X			X		
Course FE	X	X		X	X		X	X	
Year FE	X		X	X		X	X		X
Student-Year FE		X			X			X	
Student-Course FE			X			X			X

Student-clustered standard errors shown in parentheses. All models include time-variant student, teacher, and class-averaged characteristics.

Table C2: Fixed Effect Level Course Analysis on Learning Skills Index

	(1) Learning Skills Index	(2) Learning Skills Index	(3) Learning Skills Index
RM × Asian	0.072***	0.068***	0.057***
	(0.002)	(0.002)	(0.003)
RM × Black	0.068***	0.067***	0.063***
	(0.002)	(0.002)	(0.002)
RM × Latino	0.058***	0.062***	0.058***
	(0.001)	(0.001)	(0.001)
RM × White	-0.004**	-0.004**	0.000
	(0.001)	(0.001)	(0.002)
N	16838307	16819231	15977999
Student FE	X		
Course FE	X	X	
Year FE	X		X
Student-Year FE		X	
Student-Course FE			X

Student-clustered standard errors shown in parentheses. All models include time-variant student, teacher, and class-averaged characteristics.

Table C3: Race Match Effects on Individual Components with Lagged Outcome included in Model

Panel A Overall	(1) GPA	(2) Work Habit	(3) Cooperation	(4) Held Back	(5) Suspension	(6) ln(Days Absent)
Share RM	0.026***	0.033***	0.020***	-0.002	-0.000	-0.010
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.006)
Panel B By Race	**(1) GPA**	**(2) Work Habit**	**(3) Cooperation**	**(4) Held Back**	**(5) Suspension**	**(6) ln(Days Absent)**
Asian × Share RM	0.073***	0.080***	0.079***	-0.015***	-0.003	-0.174***
	(0.012)	(0.012)	(0.011)	(0.004)	(0.003)	(0.027)
Black × Share RM	0.063***	0.070***	0.063***	-0.007	-0.001	-0.069***
	(0.009)	(0.009)	(0.009)	(0.004)	(0.005)	(0.017)
Latino × Share RM	0.018**	0.029***	0.021***	0.000	-0.001	0.026*
	(0.006)	(0.006)	(0.006)	(0.003)	(0.002)	(0.011)
White × Share RM	0.011	0.012	-0.021**	0.002	0.002	-0.002
	(0.007)	(0.007)	(0.007)	(0.003)	(0.003)	(0.014)
N	1,712,241	1,712,241	1,712,239	1,712,241	1,712,241	1,712,241

Student-clustered standard errors shown in parentheses. All columns include school-grade, year, and student fixed effects. All models include time-variant student, teacher, class-averaged characteristics, and a lag of the outcome.

Table C4: Race Match Effects on Indices with Lagged Outcome included in Model

Panel A Overall	(1) Behavior Index	(2) Learning Skills Index
% RM	0.044^{***}	0.038^{***}
	(0.005)	(0.004)
Panel B By Race	**(1) Behavior Index**	**(2) Learning Skills Index**
Asian × % RM	0.155^{***}	0.120^{***}
	(0.018)	(0.014)
Black × % RM	0.109^{***}	0.085^{***}
	(0.017)	(0.011)
Latino × % RM	0.028^{**}	0.031^{***}
	(0.010)	(0.007)
White × % RM	0.003	-0.000
	(0.012)	(0.009)
N	1,712,241	1,712,241

Student-clustered standard errors shown in parentheses. All columns include school-grade, year, and student fixed effects. All models include time-variant student, teacher, and class-averaged characteristics.

Appendix D: Criteria for Marks

Los Angeles Unified School District
POLICY BULLETIN ATTACHMENT A

CRITERIA FOR MARKS

Academic Mark	A	B	C	D	Fail
Quality of Work	Demonstrates an exemplary level of understanding of content standards and tasks.	Demonstrates a thorough understanding of the content standards and tasks.	Demonstrates an understanding of the content standards and tasks.	Demonstrates a limited understanding of the content standards and tasks.	Demonstrates an inability to understand the content standards and tasks.
Interpretation and Application	Demonstrates exceptional and fluent skills in analyzing, synthesizing, and drawing inferences from observations and other data or information.	Demonstrates fluent skills in analyzing, synthesizing, and drawing inferences from observations and other data or information.	Demonstrates satisfactory skills in analyzing, synthesizing, and drawing inferences from observations and data or information.	Demonstrates a limited ability to analyze, synthesize, and draw inferences from observations and other data or information.	Demonstrates an incomplete and/or inaccurate analysis of data or information that has been collected.
Thinking and Reasoning Skills	Demonstrates an insightful and thorough use of prior knowledge and skills to create innovative ideas, products or performances in a variety of contexts.	Demonstrates an insightful use of prior knowledge and skills to create innovative ideas, products or performances in a variety of contexts.	Demonstrates use of prior knowledge and skills to create innovative ideas, products or performances in a variety of contexts.	Demonstrates limited use of prior knowledge and skills to create innovative ideas, products or performances.	Demonstrates incomplete use of prior knowledge/skills to create innovative ideas, products or performances.
Quantity of Work	Produces extra work in addition to assigned work, of both teacher-generated and self-initiated toward achieving standards for the course.	Produces extra work in addition to all assigned work, usually teacher-generated and self-initiated toward achieving standards for the course.	Produces the assigned work in achieving standards for the course.	Demonstrates a need to improve in the amount of work completed and effort expended toward achieving standards for the course.	Demonstrates no improvement of the work completed and in the effort expended toward achieving standards for the course.

WORK HABITS	E	S	U
Effort	Demonstrates exceptional determination in accomplishing tasks and mastering standards.	Demonstrates determination in accomplishing tasks and mastering standards.	Demonstrates little determination in accomplishing tasks and mastering standards.
Responsibility	Accepts complete responsibility for personal actions and demonstrates honesty, fairness, and integrity.	Accepts responsibility for personal actions and frequently demonstrates honesty, fairness, and integrity.	Accepts little responsibility for personal actions.
Attendance	Maintains excellent attendance record by consistently avoiding unnecessary absences or tardies.	Maintains a satisfactory attendance record by avoiding unnecessary absences or tardies.	Makes little effort to maintain a satisfactory attendance record; is frequently absent or tardy without excuses.
Evaluation	Makes explicit effort to examine work using both teacher-generated and self-generated criteria.	Makes effort to examine work using teacher-generated criteria.	Makes use only of teacher-generated criteria to examine work on an inconsistent basis.

COOPERATION	E	S	U
Courtesy	Maintains courteous relations with the teacher and other students and consistently works without disturbing others.	Demonstrates courteous relations with the teacher and other students and generally works without disturbing others.	Demonstrates discourteous behavior towards the teacher and other students and consistently lacks consideration for others.
Conduct	Obeys rules, respects public and personal property and actively promotes the general welfare.	Obeys rules, respects public and personal property and supports the general welfare.	Shows disregard for rules; has little respect for public and personal property and often opposes the general welfare.
Improvement	Assumes responsibility for personal improvement and rarely needs correction.	Tries to improve and usually accepts corrections in an objective manner.	Makes little attempt to improve and shows indifference or resistance to corrections.
Class Relations	Demonstrates leadership ability to work with others in a variety of situations to set and achieve goals.	Demonstrates ability to work with others in a variety of situations to set and achieve goals.	Demonstrates little ability to work with others in a variety of situations to set and achieve goals.