

ESSAYS ON THE ECONOMICS OF EDUCATION

By

William Jesse Wood

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Economics – Doctor of Philosophy

2022

## ABSTRACT

### ESSAYS ON THE ECONOMICS OF EDUCATION

By

William Jesse Wood

**Chapter 1: The Student-Teacher Race Match Effect on Noncognitive Skills.** In this chapter, I provide evidence that diversifying the labor supply of teachers to better reflect the racial distribution of students improves noncognitive outcomes for students of color without diminishing outcomes for White students. I use administrative data spanning 2007 to 2017 from 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 and behavior outcomes: GPA, work habits, cooperation, grade retention, suspensions, absences, and a data-generated noncognitive index. I mitigate the concern that race matches are endogenous by including school-grade and student fixed effects in a linear regression model. My findings indicate that students of color are expected to experience increases in GPA, work habits, and cooperation and see decreases in suspensions and 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.

**Chapter 2: The Effect of Same Race Teachers and Faculty on Student Test Scores.** This chapter estimates the impact of race matched faculty (i.e., any teacher outside of a particular student's classroom) on student test scores. While the student population rapidly continues to diversify, the diversification of the teaching corps continues to lag behind. For example, the proportion of Latino student enrollment in public schools has increased from 11 to 27 percent in just the last two decades. In contrast, share of Latino public school teachers during this same period has increased from 3

to only 9 percent (Pew Research Center, 2021). If the disparity between student and teacher racial distributions continues to grow, students of color may find it more difficult to benefit from direct student-teacher race matching (e.g., Wood, 2022). However, it may still be possible for students to benefit from same-race teachers even if they are not placed in the same classroom. Written jointly with Ijun Lai, we use administrative panel data between school years 2008-09 through 2017-18 from Los Angeles Unified School District and estimate that Latino students see positive impacts of race matched faculty. By basing this study in an area which has a large proportion of Latino students and teachers, we can fill a gap within the literature by examining the effects of race match and faculty race match on Latino students. Our findings indicate that matching Latino students to racially congruent teachers and faculty can improve math and English Language Arts test scores. Increasing the supply of Latino teachers may provide as a crucial catalyst in decreasing the achievement gaps found between Latino and white students.

**Chapter 3: Are Effective Teachers for Students with Disabilities Effective Teachers for All?** The success of many students with disabilities (SWDs) depends on access to high-quality general education teachers. Yet, most teacher value-added measures (VAMs) fail to distinguish between a teacher's effectiveness in educating students with and without disabilities. We create two VAMs: one focusing on teachers' effectiveness in improving outcomes for SWDs, and one for non-SWDs. We find top-performing teachers for non-SWDs often have relatively lower VAMs for SWDs, and that SWDs sort to teachers with lower scores in both VAMs. Overall, SWD-specific VAMs may be more suitable for identifying which teachers have a history of effectiveness with SWDs and could play a role in ensuring that students are being optimally assigned to these teachers.

Copyright by  
WILLIAM JESSE WOOD  
2022

## **ACKNOWLEDGEMENTS**

I have received a great deal of support and guidance throughout the writing of this dissertation. I would first like to thank my advisor, Scott Imberman, whose expertise was crucial in refining my research questions and methodologies. Your feedback helped to elevate the level of my work to something I am proud to have produced. I would also like to thank Katharine Strunk for her guidance throughout my studies and for repeatedly employing me as a graduate assistant. You provided me with the tools, environment, and network I needed to transform into a researcher and complete my dissertation. My gratitude also goes to my committee members, Todd Elder and Jeffrey Wooldridge, who were always able to lift my spirits in addition to their thoughtful feedback. I depended heavily on the advice of Ijun Lai, who always made me laugh when I needed it most. I would also like to thank the Los Angeles Unified School District for providing the data necessary to conduct this research. Particularly Jolene Chavira, Inocencia Cordova, Ileana Davalos, Linda Del Cueto, Kathy Hayes, Crystal Jewett, Joshua Klarin, Jonathan Lesser, and Kevon Tucker-Seeley for their assistance in obtaining and understanding the data. Thank you, Mah and Faddy, for your invaluable support, dependability, and love. Most importantly, I want to thank my wife Anna and my teeny-tiny dogs, Cleo and Lady. Blub forever.

## TABLE OF CONTENTS

<b>LIST OF TABLES . . . . .</b>	<b>ix</b>
<b>LIST OF FIGURES . . . . .</b>	<b>xi</b>
<b>KEY TO ABBREVIATIONS . . . . .</b>	<b>xiii</b>
<b>CHAPTER 1 THE STUDENT-TEACHER RACE MATCH EFFECT ON NONCOGNITIVE SKILLS . . . . .</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Theoretical mechanisms for the race match effect . . . . .	3
1.2.1 Student-side mechanisms . . . . .	3
1.2.2 Teacher-side mechanisms . . . . .	3
1.3 Empirical Strategy . . . . .	5
1.3.1 Los Angeles Unified School District Administrative Data . . . . .	5
1.3.2 Noncognitive Outcome Variables . . . . .	6
1.3.3 Descriptive Findings . . . . .	8
1.3.4 Latent Noncognitive Ability . . . . .	12
1.3.5 Models & Identification Strategy . . . . .	16
1.4 Student-Teacher Race Match Effects on Noncognitive Skills . . . . .	18
1.4.1 Individual Noncognitive Outcomes . . . . .	18
1.4.2 Noncognitive Index . . . . .	20
1.5 Course-Level Results for GPA, Work Habits, and Cooperation . . . . .	21
1.6 Robustness Checks . . . . .	27
1.6.1 Course-Level Altering Fixed Effects . . . . .	27
1.6.2 Course-Level Placebo Tests . . . . .	28
1.6.3 Lagged Outcomes . . . . .	33
1.6.4 Placebo Test using Following Year Race Match . . . . .	35
1.7 Non-linear Race Match Effect . . . . .	36
1.8 Conclusion & Implications . . . . .	42
<b>CHAPTER 2 THE EFFECT OF SAME RACE TEACHERS AND FACULTY ON STUDENT TEST SCORES . . . . .</b>	<b>43</b>
2.1 Disclaimer . . . . .	43
2.2 Introduction . . . . .	43
2.3 Theoretical Mechanisms and Relevant Literature . . . . .	45
2.3.1 Student-teacher race match effect on test scores . . . . .	45
2.3.2 Teacher peer effects, school culture, and other-teacher/staff . . . . .	47
2.4 Data . . . . .	48
2.5 Empirical Strategy . . . . .	51
2.5.1 Model . . . . .	51
2.5.2 Endogeneity of Faculty-Student Race Matching . . . . .	52

2.5.3	Endogeneity of Student-Teacher Race Matching . . . . .	53
2.6	Results . . . . .	53
2.6.1	Overall direct and faculty race match effects . . . . .	53
2.6.2	Direct and faculty race match effects by student race . . . . .	55
2.7	Robustness Checks . . . . .	57
2.7.1	Additional Controls for Student Tracking . . . . .	57
2.7.2	Placebo Test . . . . .	58
2.8	Race Match by Elementary and Middle School Grades . . . . .	60
2.8.1	Faculty Race Match Quartiles . . . . .	61
2.9	Conclusion . . . . .	63
<b>CHAPTER 3 ARE EFFECTIVE TEACHERS FOR STUDENTS WITH DISABILITIES EFFECTIVE TEACHERS FOR ALL? . . . . .</b>		<b>65</b>
3.1	Disclaimer . . . . .	65
3.2	Introduction . . . . .	65
3.3	Data . . . . .	68
3.3.1	Background and Context . . . . .	68
3.3.2	Sample and Data Requirements for Generating SWD and non-SWD VAMs . . . . .	69
3.3.3	SWD and non-SWD VAM Calculations . . . . .	73
3.4	Research Question 1: Is there evidence that a substantial portion of teachers have a relative advantage in teaching SWD vs non-SWDs and vice-versa? . . . . .	74
3.4.1	RQ1 Methods . . . . .	74
3.4.2	RQ1 Results . . . . .	76
3.5	Research Question 2: Are higher SWD/non-SWD VAMs associated with observable teacher characteristics? If so, do teachers with higher SWD VAMs share the same observable characteristics as teachers with higher non-SWD VAMs? . . . . .	80
3.5.1	RQ2 Methods . . . . .	80
3.5.2	RQ2 Results . . . . .	81
3.6	Research Question 3: Do SWDs sort into classes taught by a teacher with high (or low) SWD VAMs? . . . . .	83
3.6.1	RQ3 Methods . . . . .	83
3.6.2	RQ3 Results . . . . .	84
3.7	Research Question 4: Are SWDs sorted to teachers with relatively higher SWD VAMs compared to their non-SWD VAMs? . . . . .	85
3.7.1	RQ4 Methods . . . . .	85
3.7.2	RQ4 Results . . . . .	87
3.8	Research Question 5: How do retention rates compare for teachers with high versus low SWD VAMs? Are schools more or less likely to retain teachers with relative advantages for instructing SWDs than those with relative disadvantages? . . . . .	89
3.8.1	RQ5 Methods . . . . .	89
3.8.2	RQ5 Results . . . . .	90
3.9	Discussion . . . . .	91
<b>APPENDICES . . . . .</b>		<b>94</b>

APPENDIX A	STUDENT DESCRIPTIVE CHARACTERISTICS AND NONCOGNITIVE SKILL CORRELATIONS . . . . .	95
APPENDIX B	CRITERIA FOR MARKS IN LAUSD . . . . .	97
APPENDIX C	EXPLORATORY FACTOR ANALYSIS . . . . .	98
APPENDIX D	RACE MATCHED EFFECTS FOR NONCOGNITIVE COMPONENTS	99
APPENDIX E	SUMMARY STATISTICS, RESULTS, AND PLACEBO TEST FOR COURSE LEVEL ANALYSIS . . . . .	102
APPENDIX F	ROBUSTNESS CHECKS FOR STUDENT-TEACHER RACIAL SORTING WITHIN SCHOOLS . . . . .	108
APPENDIX G	ROBUSTNESS CHECK FOR THRESHOLD LEVELS . . . . .	111
APPENDIX H	SWD SORTING BY DISABILITY TYPE . . . . .	113
APPENDIX I	RELATIONSHIP BETWEEN VAM/RELATIVE ADVANTAGE AND SWITCHING SCHOOLS WITHIN LAUSD . . . . .	115
<b>BIBLIOGRAPHY</b>	. . . . .	<b>116</b>



## LIST OF TABLES

Table 1.1	Student Descriptive Characteristics . . . . .	6
Table 1.2	Cross-correlations of Noncognitive and Cognitive Measures . . . . .	15
Table 1.3	Student-Teacher Race Match Effects on Noncognitive Outcomes . . . . .	19
Table 1.4	Race Match Effects including Lagged Outcome . . . . .	34
Table 2.1	Student & Teacher Summary Statistics by Race . . . . .	50
Table 2.2	Overall Direct and Faculty Race Match Effects on Test Scores . . . . .	54
Table 2.3	Direct and Faculty Race Match Effects on Test Scores by Student Race . . . . .	56
Table 2.4	Race match effects with additional controls for student tracking . . . . .	58
Table 2.5	Race Match Effects on Test Scores by Elementary/Middle School for Latino Students . . . . .	60
Table 3.1	Student and Teacher Descriptive Characteristics . . . . .	71
Table 3.2	Share of Teachers by Descriptive Characteristics, Varying Minimum Threshold of Students . . . . .	73
Table 3.3	Teacher Distribution of SWD and Non-SWD VAMs by Quintile . . . . .	78
Table 3.4	Association between Teacher Characteristics and SWD and non-SWD VAMs . . . . .	82
Table 3.5	Association between Student Disability Status and Teacher VAMs . . . . .	84
Table 3.6	Association between Student Disability Status and Teacher VAMs . . . . .	86
Table 3.7	Association between Student Disability Status and Teacher Relative Advantage . . . . .	88
Table 3.8	VAMs and Relative Advantage Association with Leaving LAUSD . . . . .	91
Table A.1	Average of Teacher Characteristics by Student Race . . . . .	95
Table A.2	Summary of Share Race Match and Student Noncognitive Skills Measures . . . . .	96
Table A.3	Cross-correlations of Noncognitive Components . . . . .	96

Table C.1	Factor Loadings and Eigenvalues for Exploratory Factor Analysis . . . . .	98
Table E.1	Percent of Students with Direct Race Match by Subject . . . . .	102
Table E.2	Percent of Teachers by Subject across Teacher Race . . . . .	103
Table E.3	Direct Race Match Effect at Course Level . . . . .	104
Table G.1	Association between Student Disability Status and Teacher VAMs by Thresh- old Sample . . . . .	112
Table H.1	Association between Student Disability Type and Teacher VAMs . . . . .	113
Table H.2	Association between Student Disability Type and Teacher Relative Advantage .	114
Table I.1	VAMs and Relative Advantage Association with Switching Schools within LAUSD	115

## LIST OF FIGURES

Figure 1.1	National- and LAUSD-Level Teacher and Student Racial Distributions . . . . .	8
Figure 1.2	Average Percent (%) Race Match in LAUSD . . . . .	9
Figure 1.3	Summary of GPA, Work Habits, and Cooperation per Year by Race . . . . .	10
Figure 1.4	Summary of Held Back, Suspensions, and Days Absent per Year by Race . . . . .	11
Figure 1.5	Component Loading across Noncognitive and Cognitive Indices . . . . .	14
Figure 1.6	Race Match Effect on Learning Skills and Behavior . . . . .	20
Figure 1.7	Direct Race Match Effect at Course Level . . . . .	24
Figure 1.8	Correlation of Direct Race Matches across Courses by Student Race . . . . .	30
Figure 1.9	Race Match Effects by Placebo Race Match Status (ELA) . . . . .	31
Figure 1.10	Placebo Test using following Year Share Race Match . . . . .	36
Figure 1.11	Race Match Effects by Quintile of Share RM . . . . .	38
Figure 1.12	Race Match Effects on Noncognitive Index by Quintile of Share RM . . . . .	41
Figure 2.1	Placebo Test using following Year Race Match Components . . . . .	59
Figure 2.2	Math Test Scores: Effects by Quintile of Share Faculty Race Match . . . . .	62
Figure 2.3	ELA Test Scores: Effects by Quintile of Share Faculty Race Match . . . . .	63
Figure 3.1	Comparison of Teacher SWD and non-SWD VAMs . . . . .	77
Figure 3.2	DVAM and Random DVAM Distributions . . . . .	79
Figure B.1	Rubric for Marks for Achievement, Work Habits, and Cooperation . . . . .	97
Figure D.1	Race Match Effect on Standardized Noncognitive Measures . . . . .	99
Figure D.2	Race Match Effect on the Probability of Noncognitive Measures . . . . .	100
Figure D.3	Race Match Effect on ln(Days Absent) . . . . .	101

Figure E.1	Race Match Effects by Placebo Race Match Status (Math) . . . . .	105
Figure F.1	Race Match Effect with Lag for Standardized Noncognitive Outcomes . . . . .	108
Figure F.2	Race Match Effect with Lag for Probability Noncognitive Outcomes . . . . .	109
Figure F.3	Race Match Effect with Lag for ln(Days Absent) . . . . .	110
Figure G.1	DVAM and Random DVAM Distributions . . . . .	111

## KEY TO ABBREVIATIONS

**DVAM** Difference in Value-Added Measures

**EFA** Exploratory Factor Analysis

**EL** English Learner

**ELA** English language arts

**FE** Fixed Effect

**FRM** Faculty Race Match

**GET** General Education Teacher

**GPA** Grade Point Average

**LAUSD** Los Angeles Unified School District

**LS Imp.** Language and Speech Impairment

**Noncog.** Noncognitive

**PCA** Principal Component Analysis

**Rel. Adv.** Relative Advantage

**RM** Race Match

**RQ** Research Question

**sd** Standard Deviation

**SLD** Specific Learning Disability

**SWD** Student with Disability

**VAM** Value-Added Measure

## **CHAPTER 1**

### **THE STUDENT-TEACHER RACE MATCH EFFECT ON NONCOGNITIVE SKILLS**

#### **1.1 Introduction**

There continue to be large racial gaps across education outcomes in the United States. 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 learning skills and behavior also exist between White and non-White students, even when controlling for environmental factors (Elder and Zhou, 2021; Hashim et al., 2018). Gaps in these non-test outcomes, referred to as “noncognitive” skills, are just as concerning given their correlation with long-run outcomes, such as high school graduation, college enrollment, college major choice, employment, and wages (e.g., Price, 2010; Heckman et al., 2006, 2013; Jackson, 2018).

Prior research suggests that one potential avenue to close educational gaps could be through matching students with teachers of the same race. For example, research has established that test scores for students of color improve when students are matched to teachers who share the same racial background (e.g., Dee, 2004; Egalite et al., 2015; Joshi et al., 2018). A rapidly growing body of research suggests that race matching can also significantly affect non-test score outcomes. For example, Grissom et al. (2017) find that race matched Black and Latino students are more likely to be recommended for enrollment in gifted education programs. Access to the additional resources, teachers, and peers within these programs could improve a student’s noncognitive skills. Harbatkin (2021) found that race matched Black students receive higher grade point averages (GPA) than non-race matched Black students. Race matched students are also significantly less likely to face disciplinary actions compared to their peers who were not race matched (Lindsay and Hart, 2017; Shirrell et al., 2021). However, relatively little remains known about other education outcomes such as attendance and social-emotional skills like work habits and cooperation.

This paper contributes to our growing understanding of race match and student noncognitive

skills in three critical ways. First, I examine how student-teacher race matching impacts multiple non-test score outcomes. Using data from the Los Angeles Unified School District (LAUSD), I study the relationship between student-teacher race match and suspensions, absences, grade retention, grade point average, work habits, and cooperation. I find increases in GPA, marks for work habits, 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.

It is likely that a student's cognitive abilities may influence these outcomes. For example, a student's GPA, usually calculated by their in-class assignments and test scores, may be affected by their cognitive skills. As a second contribution, I employ a factor analysis to generate a noncognitive index, which proxies for a student's ability to manage their time, exercise self-control, cooperate with classmates, resolve conflicts appropriately, and externalizing behaviors such as aggression and disruptiveness. By design, this index is orthogonal to test scores, a commonly used proxy for a student's cognitive abilities. I find that race matched secondary school (i.e., grades 6 through 12) students of color experience increases in this noncognitive measure.

As a third contribution, this study broadens the scope of race match studies beyond traditionally studied races and settings. While there is a growing body of research that descriptively links race matched Latino students and educational outcomes, such as taking advanced courses (e.g., Kettler and Hurst, 2017), there are few causal studies on this topic. Instead, previous causal studies have focused on the race match effects for Black and White students,<sup>1</sup> even though Latino students make up a quarter of the student population (NCES, 2017) and face large gaps in educational outcomes compared to White students. Additionally, most of the data used in these studies is at the state or national level, which does not allow researchers to answer whether race match effects exist in diverse urban areas, where many students of color are located (Pew Research Center, 2018). By leveraging the racial composition and size of LAUSD, I can explore how race match effects differ in an urban setting and shed light on previously understudied groups of students. In this setting, I

---

<sup>1</sup>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.

find large and significant increases in several non-test score outcomes, as well as the noncognitive index, for Asian, Black, and Latino students.

## **1.2 Theoretical mechanisms for the race match effect**

### **1.2.1 Student-side mechanisms**

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. 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 underperforming in their classes. When students are placed with a teacher of the same race, they may be less likely to worry about teachers using negative stereotypes to judge their actions. Instead, students of color may feel more comfortable asking questions and seeking help in race matched classrooms. Correspondingly, students may improve their academic effort, classroom cooperation, and grades.

Race match may also influence student-side changes by providing students of color with a positive role model. For example, students of color may be inspired by teachers' own past and trajectory. Past studies have linked positive role models to increased high school graduation as well as social-emotional outcomes such as confidence and behavior (e.g., Gershenson et al., 2018; Goings and Bianco, 2016).

### **1.2.2 Teacher-side mechanisms**

Studies suggest several teacher-side mechanisms for race match effects. To begin, White teachers may have implicit biases against students of color, which may contribute to how teachers perceive students' actions. For example, psychological experiments have shown that the same exact misbehavior is perceived as more disruptive and deserving of harsher punishment when the student is Black (e.g., Okonofua et al., 2016). Additionally, a recent study by Chin et al. (2020)



has linked higher suspension rates for Black students with higher teacher self-reports of implicit bias—suggesting that biases may play an important role in disproportionate discipline rates for students of color (e.g., Shi and Zhu, 2021; Liu et al., 2022; Skiba et al., 2011; Rodriguez, 2020). Studies have also shown that Black teachers have lower levels of implicit bias (e.g., Chin et al., 2020) and that Black and Latino students were less likely to be seen as disruptive when they were race matched to their teachers (e.g., Dee, 2005).

Teacher expectations are another teacher-side mechanism that could explain race match differences. Previous studies have long documented that White teachers tend to have lower academic expectations for students of color (e.g., Diamond et al., 2004), both in comparison to White students as well as their actual abilities (Fox, 2015; Gershenson et al., 2016; Ready and Wright, 2011; Tenenbaum and Ruck, 2007). Lower expectations often translate into easier classroom assignments for and lower effort from students of color (e.g., Ferguson, 2003; Grissom et al., 2015). 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).

Race matched teachers may also benefit students of color through the increased likelihood of developing culturally relevant curriculum and classroom strategies that meet students' academic as well as social and emotional needs (Gay, 2000; Milner, 2011; Saft and Pianta, 2001; Ladson-Billings, 1995). For example, previous studies documented how Latino teachers have brought their students' cultural and linguistic backgrounds into the classroom to create engaging and caring educational environments (Antrop-González and De Jesus, 2006; Newcomer, 2018). Furthermore, a recent descriptive study by Lopez (2016) has shown a correlation between Latino teachers using culturally responsive teaching practices and higher reading achievement in Latino students. More generally, race matched students were more likely to report feeling motivated, engaged, and cared for than when they were paired with a White teacher (Cherng and Halpin, 2016; Egalite et al., 2015).

## 1.3 Empirical Strategy

### 1.3.1 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. 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 grade point averages and suspensions, teachers within this district assess students based on other key noncognitive skills: work habits and cooperation.

For students, LAUSD records race,<sup>2</sup> 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,<sup>3</sup> 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 noncognitive skills, I collapse the data to the student-year level. Table 1.1 summarizes each student characteristic and the average share of teacher race that shares a classroom with students by student race.<sup>4</sup> The majority of students of color are FRL eligible, a proxy measure for a student's socioeconomic status, with 61.2%, 71.5%, and 83.3% of Asian, Black, and Latino students being eligible for free- or reduced-price lunches, respectively. As a reference, the average for all public-school students in California in 2017 was 59.4% (NCES, 2017). On average, Asian (13.3%) and Latino (24.5%) students are classified as English Learners at higher rates than the

---

<sup>2</sup>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.

<sup>3</sup>Since years of experience is top-coded at 10 years, this study groups experience into four categories: novice (less than or equal to 2 years), early-career (3 through 5 years), mid-career (6 through 7 years), and late-career (more than 8 years).

<sup>4</sup>Table A.1 in Appendix A summarizes the average of each teacher characteristic by student race.

Black (0.1%) and White (6.3%) students in the sample. Black (6.8%) students are identified as students with disabilities more often than the remaining races in the sample (4.9% overall). Since these characteristics may influence a student's noncognitive outcomes, they are controlled for in the models detailed in Section 1.3.5 below.

**Table 1.1** Student Descriptive Characteristics

	Asian	Black	Latino	White	Total
Female	0.481 (0.500)	0.501 (0.500)	0.490 (0.500)	0.474 (0.499)	0.489 (0.500)
FRL	0.612 (0.487)	0.715 (0.451)	0.833 (0.373)	0.403 (0.490)	0.770 (0.421)
EL	0.133 (0.339)	0.010 (0.098)	0.245 (0.430)	0.063 (0.243)	0.199 (0.399)
SWD	0.030 (0.171)	0.068 (0.252)	0.048 (0.214)	0.046 (0.209)	0.049 (0.215)
Total Tch.	5.878 (1.526)	5.987 (1.948)	5.848 (1.774)	5.862 (1.622)	5.865 (1.767)
Share Asian Tch.	0.140 (0.158)	0.085 (0.132)	0.089 (0.134)	0.099 (0.140)	0.092 (0.137)
Share Black Tch.	0.070 (0.123)	0.257 (0.265)	0.098 (0.164)	0.063 (0.119)	0.108 (0.177)
Share Latino Tch.	0.196 (0.198)	0.170 (0.187)	0.319 (0.251)	0.174 (0.182)	0.285 (0.245)
Share White Tch.	0.550 (0.247)	0.435 (0.271)	0.435 (0.264)	0.620 (0.236)	0.458 (0.267)
N	112,037	245,796	2,016,692	212,427	2,674,791
Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. The mean characteristics are displayed with the standard deviations in parentheses. <i>Share [Race] Tch.</i> represents the average share of [race] teachers that a student has in a given year; <i>FRL</i> = free- or reduced-price lunch eligible; <i>EL</i> = English learner; <i>SWD</i> = student with disability.					

### 1.3.2 Noncognitive Outcome Variables

Each year, a student is assigned a mark for achievement (intended to capture quality of work, reasoning skills, and ability to analyze information), work habits (intended to capture effort,

responsibility, attendance, and self-evaluation), and cooperation (intended to capture courtesy, conduct, personal improvement, and class relations) for every course (e.g., math, science, social studies, etc.) they are enrolled. A detailed rubric for how each mark for achievement, work habits, and cooperation are graded within LAUSD is provided in Appendix B. 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.<sup>5</sup> 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." The dataset includes the attendance record (*Days Absent*) for each student at the academic-year level. Suspension data for each student is also reported at the academic-year level, and I use this data to generate an indicator variable for whether a student was ever suspended during the school year (*Suspension*). The panel nature of the dataset allows for the identification of whether a student is retained into the same grade the following year, which I utilize to create an indicator for whether a student was held back (*Held Back*).

Each of these components may be influenced by a student's latent noncognitive abilities. When calculating a student's GPA, teachers often incorporate student behavior, attitude, work habits, and effort as factors (Guskey, 2018; Howley et al., 2000). Duckworth et al. (2007) find that students with fewer absences, suspensions, and grade retentions are associated with higher levels of conscientiousness, which they characterize as individuals who are "thorough, careful, reliable, organized, industrious, and self-controlled." Barbaranelli et al. (2003) find that low levels of conscientiousness, regardless if rated by a child's mother or their teacher, have been linked to higher levels of externalizing behaviors (e.g., aggression and impulsivity). In the next section, I describe how these noncognitive measures differ across student race and race match status within LAUSD.

---

<sup>5</sup>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.

### 1.3.3 Descriptive Findings

Figure 1.1 shows the larger share of Latino teachers and students in LAUSD than found on average across 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 75.4% for teachers and students, respectively. The larger share of the Latino population within LAUSD is critical in allowing for a focus on a traditionally understudied group within the race match literature. Another takeaway from Figure 1.1 is the disparity between the student and teacher racial distributions across both panels. For example, nationally, over half (50.5%) of students are of color. However, more than four-fifths (81.7%) of teachers are White. In LAUSD, 92.1% of students are of color, but White teachers still make up the plurality of the teaching workforce at 44.6%.

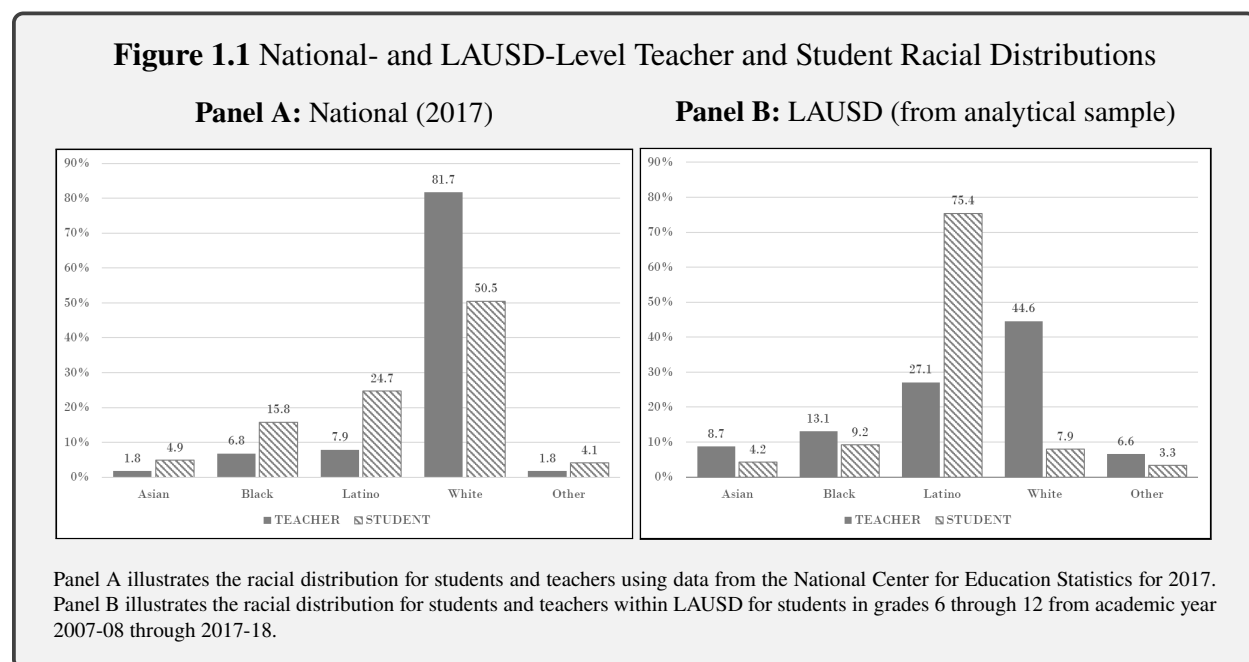
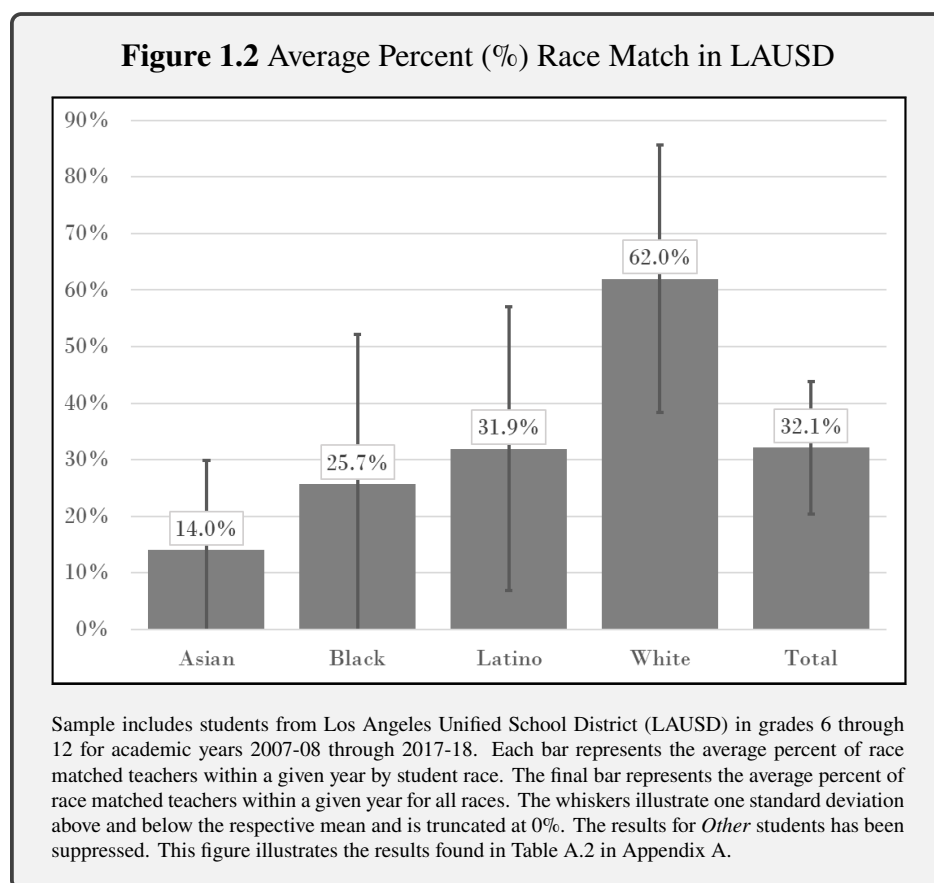


Figure 1.2 illustrates the differences in the treatment variable (*Share RM*) across student race in LAUSD. *Share RM* is the share of race matched teachers a student has in a given year.<sup>6</sup> On average,

<sup>6</sup>Presumably, the amount of time a student and teacher are together influences the impact that teacher may have

students are race matched to 32.1% of their teachers in a given year.<sup>7</sup> Asian, Black, and Latino students are race matched to 14.0, 25.7, and 31.9 percent of their teachers, respectively. Despite making up a relatively smaller share of the student composition, White students are race matched to 62.0% of their teachers. This higher rate of race matching for White students is most likely due to a combination of factors. As seen in the previous figure, White teachers make up the plurality in LAUSD. Additionally, White students tend to attend schools taught by mostly White teachers (NCES, 2020). These descriptive findings suggest that White students may have an inequitable opportunity to receive the potential benefits of race matched instruction as compared to students of color.



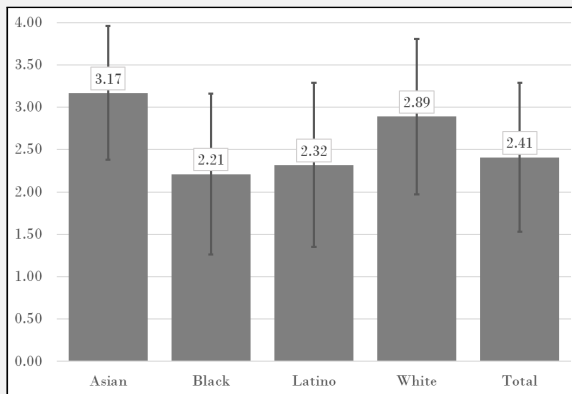
on a student's noncognitive skills. To account for this, *Share RM* is weighted by the number of times a teacher and student share a classroom within an academic year. That is,  $ShareRM_{it} = \frac{\sum RM_{ict}}{Tch_{ict}}$  where  $RM_{ict}$  is an indicator for whether a student ( $i$ ) and teacher share the same race in a given course/subject ( $c$ ; e.g., math, science, etc.) and year ( $t$ ) and  $Tch_{ict}$  is the total number of teacher-course combinations a student has in that same year.

<sup>7</sup>See Table A.2 in Appendix A for the table analogue of Figure 1.2 and a summary of each noncognitive skill by student race.

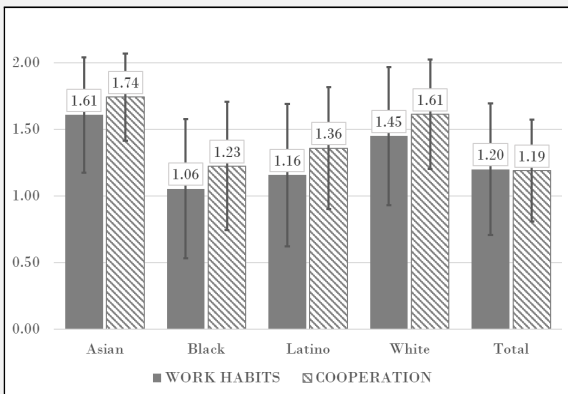
Figures 1.3 and 1.4 summarize the noncognitive components by race and illustrates the gaps found in various noncognitive outcomes between Asian/White students and Black/Latino students. On average, Asian and White students have higher GPAs, marks for work habits, and marks for cooperation than Black and Latino students. Asian and White students are also less likely to be held back, suspended, and miss fewer days. These gaps in noncognitive skills are consistent with the findings of existing literature. 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 any existing race match effect, I control for factors correlated with race matching and noncognitive outcomes, as explained in Section 1.3.5.

**Figure 1.3 Summary of GPA, Work Habits, and Cooperation per Year by Race**

**Panel A: Average GPA**



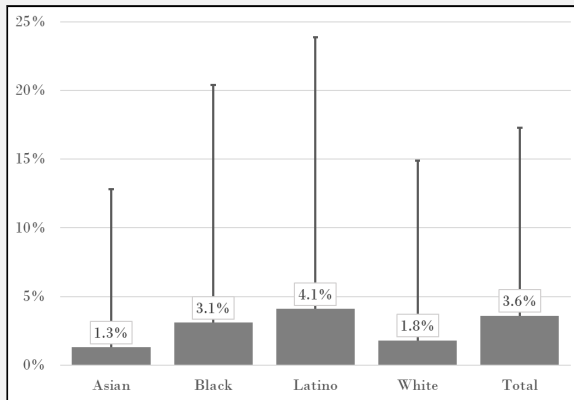
**Panel B: Average Work Habits and Cooperation**



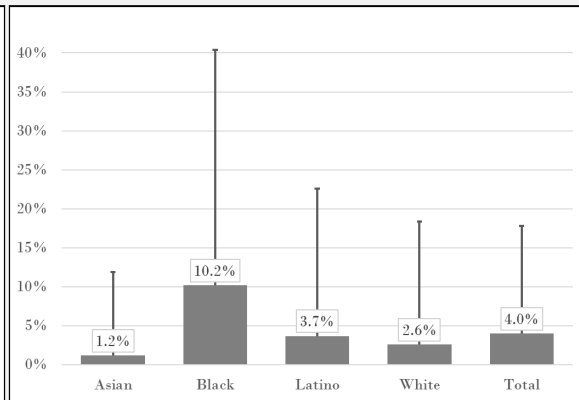
Sample includes students from Los Angeles Unified School District (LAUSD) in grades 6 through 12 for academic years 2007-08 through 2017-18. Each bar represents the mean for the given outcome within an academic year. Standard deviations are represented by the black whiskers. *Other* race results not shown. Table analogue found in Table A.2 in Appendix A.

**Figure 1.4** Summary of Held Back, Suspensions, and Days Absent per Year by Race

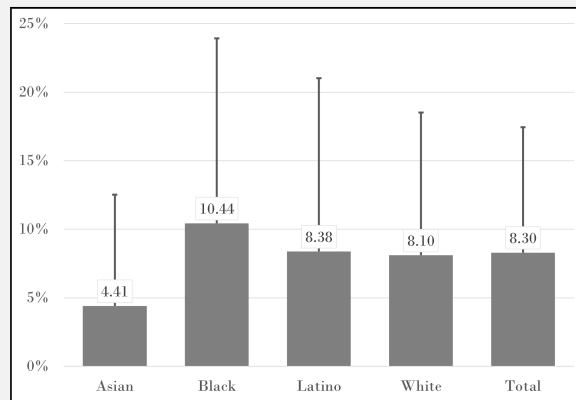
**Panel A:** Average Percent (%) Held Back



**Panel B:** Average Percent (%) Suspensions



**Panel C:** Average Percent (%) Days Absent



Sample includes students from Los Angeles Unified School District (LAUSD) in grades 6 through 12 for academic years 2007-08 through 2017-18. Each bar represents the mean for the given outcome within an academic year. Standard deviations are represented by the black whiskers. *Other* race results not shown. Table analogue found in Table A.2 in Appendix A.



### 1.3.4 Latent Noncognitive Ability

A student's cognitive ability may influence the noncognitive outcomes used in this paper. Using 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. On the other hand, teachers may assign value to a student's behavior, attitude, work habits, and effort when assigning grades (Guskey, 2018; Howley et al., 2000). To parse out as much information as possible, and because a student's noncognitive ability is difficult to directly measure, I use a latent variable framework and define a proxy for a student's noncognitive ability. Specifically, using exploratory factor analysis, I construct a *noncognitive index* to measure a student's ability to use prior knowledge to accomplish tasks, evaluate their own work, cooperate with others, and externalizing behaviors such as aggression, disruptiveness, or impulsiveness.

Exploratory factor analysis (EFA)<sup>8</sup> utilizes the shared (or common) variation across each pre-selected component (e.g., GPA, suspensions, math test scores, etc.) to calculate a proxy (i.e., *factor*) for the latent noncognitive ability. More generally, for a given  $p$  components and  $n$  observations, let  $\mathbf{X}_{p \times n} = \mathbf{L}_{p \times k} \mathbf{F}_{k \times n} + \mathbf{U}_{p \times n}$ , where  $\mathbf{X}$  is the normalized value of component  $p$ ,  $\mathbf{F}$  contains the factor scores for the  $k < p$  factors,  $\mathbf{L}$  describes how each  $p$  component loads onto each  $k$  factor, and  $\mathbf{U}$  captures the unique variation (i.e., *uniqueness*) of each component that is not shared across all other components.  $\mathbf{L}$  is estimated via eigenvector decomposition by solving for  $\text{Cov}(\mathbf{X}) = \text{Cov}(\mathbf{LF} + \mathbf{U})$ .<sup>9</sup> This paper relies on the Bartlett method (Bartlett, 1937), which minimizes the uniqueness (or maximizes the shared variation) of each component to

---

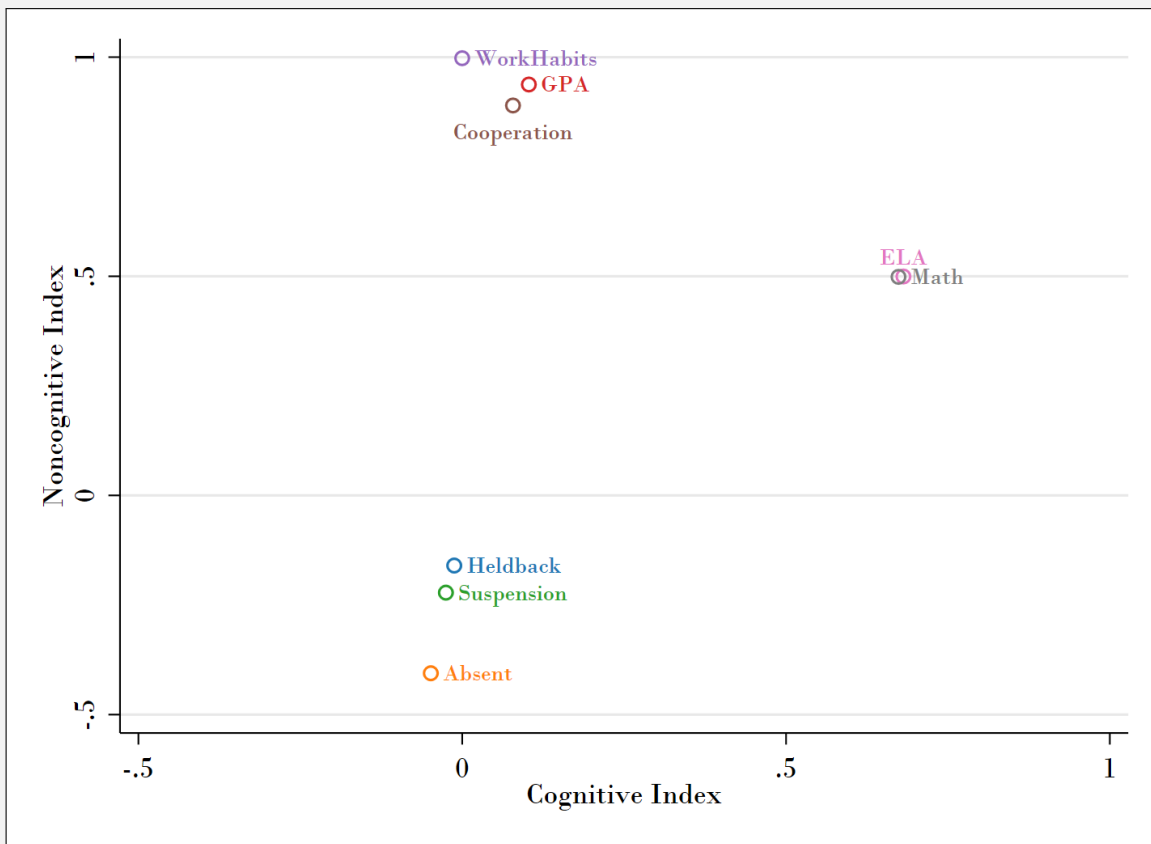
<sup>8</sup>EFA and principal component analysis (PCA) both rely on the variance structure of existing data to generate proxies for underlying latent variables. However, these two methods are fundamentally different. EFA only utilizes the common variance across each component; PCA utilizes the maximal amount of variation across each component. In EFA, the analyst has an underlying theory as to which factors influence the components (e.g., noncognitive and cognitive); PCA is an agnostic approach and does not rely on an underlying theory. See Suhr (2005) for a detailed comparison of these two methods.

<sup>9</sup>Multiple solutions exist for  $\mathbf{L}$  depending on the factor scoring method and rotation. This paper utilizes a varimax rotation, which is an orthogonal rotation method (i.e., ensures the factors are uncorrelated) that minimizes the number of components with large factor loadings on each factor. The results when using other orthogonal rotation methods, quartimax (minimizes the number of factors needed to explain the variation found in each component) and equamax (combination of quartimax and varimax), are similar to those of the varimax rotation method and are available upon request.

determine the number of orthogonal factors (i.e.,  $k$ ) and estimates their factor scores (i.e., Bartlett scores). The Bartlett method also results in unbiased estimates of the true factor scores due to the utilization of maximum likelihood estimation. Then, for the  $k$  orthogonal estimated factors,  $\hat{\mathbf{F}} = (\hat{\mathbf{L}}'\hat{\mathbf{\Psi}}^{-1}\hat{\mathbf{L}})^{-1}\hat{\mathbf{L}}'\hat{\mathbf{\Psi}}^{-1}\mathbf{X}$ , where  $\hat{\mathbf{L}}$  contains the estimated loading scores for each factor-component pairing, and  $\hat{\mathbf{\Psi}}_{p \times p}$  is a diagonal matrix whose elements are equal to the estimated uniqueness scores for each component. In the context of this paper, by observing the eigenvalues associated with each of the resulting  $k$  factors and how each of the  $p$  components *load* onto each factor, I define a proxy for a student's noncognitive skills that is orthogonal to that student's cognitive ability.

Figure 1.5 illustrates the component loadings across two factors. GPA, Work Habits, and Cooperation for a group and contribute the most towards factor 1 (eigenvalue = 2.97), which I call the *noncognitive index* (meant to measure effort, cooperation, and externalizing behaviors). Similarly, math and ELA test scores highly load onto factor 2 (eigenvalue = 1.34), which I call the *cognitive index*. The figure also shows that Heldback, Suspensions, and Absent either contribute negatively (noncognitive) or do not affect (cognitive) the factors. By design, the noncognitive and cognitive indices are orthogonal, which can be seen by the low correlation between the two measures provided in Table 1.2. Each of these measures is also normalized across students at the grade-year level.

**Figure 1.5** Component Loading across Noncognitive and Cognitive Indices



Sample includes students from Los Angeles Unified School District (LAUSD) in grades 6 through 12 for academic years 2007-08 through 2017-18 with math and ELA test scores. Observations are at the student-year level. Figure shows the component loadings post varimax rotation across the noncognitive index (eigenvalue = 2.97) and the cognitive index (eigenvalue = 1.34). See Table C.1 in Appendix C for table analogue.

**Table 1.2** Cross-correlations of Noncognitive and Cognitive Measures

	Noncognitive	Cognitive	GPA	Work Habits	Cooperation	Held Back	Absent	Suspension	Math Test	ELA Test
Noncognitive	1.000									
Cognitive	-0.004	1.000								
GPA	0.923	0.134	1.000							
Work Habits	0.982	0.010	0.931	1.000						
Cooperation	0.872	0.101	0.836	0.888	1.000					
Held Back	-0.160	-0.013	-0.130	-0.126	-0.118	1.000				
Absent	-0.406	-0.054	-0.412	-0.401	-0.369	0.122	1.000			
Suspension	-0.222	-0.028	-0.186	-0.190	-0.220	0.017	0.136	1.000		
Math Test	0.497	0.769	0.549	0.502	0.494	-0.089	-0.249	-0.115	1.000	
ELA Test	0.497	0.778	0.545	0.502	0.509	-0.082	-0.213	-0.129	0.708	1.000

Sample includes students from Los Angeles Unified School District (LAUSD) in grades 6 through 12 for academic years 2007-08 through 2017-18 with math and ELA test scores. Observations are at the student-year level. Using an exploratory factor analysis, the noncognitive index (eigenvalue = 2.97) and the cognitive index (eigenvalue = 1.34) are the factor scores using the Bartlett method and a varimax rotation.

### 1.3.5 Models & Identification Strategy

I estimate the impact that student-teacher race matching has on noncognitive measures by using the following model:

$$Y_{it} = \beta_0 + \beta_{RM}(ShareRM_{it}) + \mathbf{X}_{it}\mathbf{\Gamma}_1 + \bar{\mathbf{Z}}_{it}\mathbf{\Gamma}_2 + \bar{\mathbf{X}}_{it}\mathbf{\Gamma}_3 + \theta_i + \theta_{gs} + \theta_t + \epsilon_{igst}, \quad (1.1)$$

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 with  $\beta_{RM}$  being the coefficient of interest. While  $\beta_{RM}$  is technically the impact in the *share* of race matched teachers on student outcomes, if the model is correctly specified, it can be interpreted as the effect on noncognitive outcomes for race matched students. I test the robustness of the model specification in Section 1.6.  $\mathbf{X}$  is a vector of time-variant student characteristics (i.e., free- or reduced-price lunch eligibility, English language learner status, and student with disability status),  $\bar{\mathbf{Z}}$  is a vector of averaged teacher characteristics,<sup>10</sup>  $\bar{\mathbf{X}}$  is a vector of classroom characteristics,<sup>11</sup>  $\theta_i$ ,  $\theta_{gs}$ , and  $\theta_t$  are student, school-grade, and year fixed effects, respectively.  $\epsilon$  is a normally distributed error term. I repeat this analysis by student race using:

$$Y_{it} = \beta_0 + \beta_{RM}(\mathbb{1}(Race_i) * ShareRM_{it}) + \mathbf{X}_{it}\mathbf{\Gamma}_1 + \bar{\mathbf{Z}}_{it}\mathbf{\Gamma}_2 + \bar{\mathbf{X}}_{it}\mathbf{\Gamma}_3 + \theta_i + \theta_{gs} + \theta_t + \epsilon_{igst}, \quad (1.2)$$

where each component is identical to Equation 1.1 with the exception that  $ShareRM$  is now multiplied by an indicator variable,  $\mathbb{1}(Race)$ , which represents the student's race.

Also, for the components with data at the course-level (i.e., GPA, Work Habits, and Cooperation), I follow Harbatkin (2021) and estimate the following:

$$Y_{ict} = \beta_0 + \beta_{RM}(\mathbb{1}(Race_i) * RM_{ict}) + \mathbf{X}_{it}\mathbf{\Gamma}_1 + \mathbf{Z}_{ict}\mathbf{\Gamma}_2 + \bar{\mathbf{X}}_{ict}\mathbf{\Gamma}_3 + \theta_{gs} + \theta_i + \theta_t + \theta_c + \epsilon_{igst}, \quad (1.3)$$

<sup>10</sup>See Table A.1 for the full list of teacher controls.

<sup>11</sup>In addition to the listed student controls, the classroom characteristics also include the average of each student race and gender.

where  $RM$  is a binary variable indicating the student is race matched within a the course ( $c \in ELA, Math, History, Science, Health$ ).  $\bar{X}_{ict}$  and  $\bar{Z}_{ict}$  are the averaged classroom and teacher characteristics at the course-year level for student  $i$ .  $\theta_c$  are course fixed effects.

Empirically, students and teachers sort by race across schools (NCES, 2020), which threatens the identification of the race match effect if any unobservable (student or teacher) characteristic related to this phenomenon is correlated with noncognitive outcomes. Accounting for the racial sorting across schools may not be sufficient. For example, students (or their parents) may request a change in an assigned non-race matched teacher for a race matched teacher, which may also cause an omitted variable bias to the race match estimate. To account for these concerns, I include grade-school fixed effects ( $\theta_{gs}$ ), which utilize the *within* grade-school variation of  $ShareRM$ . Under the plausible assumptions that students do not choose to fail or skip grades based on the racial distribution of teachers within grade-schools and that teachers do not choose which grade-schools to teach based on the racial distribution of students, the grade-school fixed effects mitigate the concern for racial sorting bias between students and teachers.

Also included within the model are student fixed effects ( $\theta_i$ ), which control for any time invariant unobserved student characteristics that may lead to racial sorting and are correlated with higher noncognitive abilities. The inclusion of student fixed effects exploits the variation found in any deviations in the share of race matched teachers from a student's historical mean. In other words, the race match effect is estimated by analyzing the change in noncognitive outcomes associated with the change in the share of race matched teachers a student experiences over time. In Section 1.6, I test the identifying assumptions by altering the level of fixed effects, including lagged outcome variables, and utilizing a placebo test to check the robustness of the estimated race match effect.

However, the race match effect would still be biased if, across all years and students, teachers favorably report noncognitive outcomes for race matched students. In this case, the race match effect could no longer be interpreted as a change in the student's noncognitive skill, but rather a change in the *reporting*. Even under this scenario, it should be noted that students of color may still benefit. If more favorable grading is perceived by the student as a positive affirmation from

their teacher, that student may improve their noncognitive outcome (e.g., behavior, effort, etc.). In a meta-analysis, Wu et al. (2021) provide examples of several studies which find that teacher affirmation improves educational outcomes, especially for disadvantaged groups (e.g, students of color, female, English learners).

## 1.4 Student-Teacher Race Match Effects on Noncognitive Skills

### 1.4.1 Individual Noncognitive Outcomes

Table 1.3 shows the effects of student-teacher race matching on the individual noncognitive measures.<sup>12</sup> Panel A reports the results using Equation 1.1 and shows positive overall race match effects for GPA, marks for work habits, and marks for cooperation. Since *GPA*, *Work Habits*, and *Cooperation* are standardized at the grade-year level, their coefficients are interpreted as a change in standard deviations (sd). The race match effect is expected to increase each of these components by 0.029 sd, 0.037 sd, and 0.024 sd, respectively. This closely aligns with an existing finding for the race match effect on GPA of 0.023 sd (Harbatkin, 2021). The coefficients on *Held Back* and *Suspension* are interpreted as changes in probability. Column (4) of Table 1.3 indicates that race matched students are *more likely* to be held back a grade than non-race matched students. At first, this may seem like an undesirable outcome. However, at least in the short-run, some students may be able to benefit from grade retention (Huddleston, 2014). Race matched teachers may be able to better identify the specific needs for students and determine that the student could truly benefit from being held back. I do not find an overall race match effect on the probability of being suspended. Column (6) shows the estimated race match effect on a measure of student absenteeism. The outcome is a log transformation of the number of days a student is absent, which allows for the coefficient to be interpreted as the change in percent. Overall, I find the estimated race match effect to lower the number of missed school days by 2.1 percent.

Panel B of Table 1.3 reports the estimated race match effects by student race using Equation 1.2. For GPA, work habits, and cooperation, the race match effect similarly aligns to the sign of

---

<sup>12</sup>The figure analogues for Table 1.3 are found in Appendix D.

**Table 1.3** Student-Teacher Race Match Effects on Noncognitive Outcomes

Panel A Overall	(1) GPA	(2) Work Habits	(3) Cooperation	(4) Held Back	(5) Suspension	(6) ln(Days Absent)	(7) Noncog. Index
Share RM	0.024*** (0.002)	0.032*** (0.002)	0.019*** (0.002)	0.002** (0.001)	0.000 (0.001)	-0.021*** (0.004)	0.045*** (0.005)
Panel B By Race	(1) GPA	(2) Work Habits	(3) Cooperation	(4) Held Back	(5) Suspension	(6) ln(Days Absent)	(7) Noncog. Index
Asian × Share RM	0.106*** (0.009)	0.096*** (0.009)	0.075*** (0.008)	-0.011*** (0.002)	-0.005* (0.003)	-0.081*** (0.018)	0.140*** (0.018)
Black × Share RM	0.056*** (0.007)	0.057*** (0.007)	0.066*** (0.007)	-0.001 (0.002)	0.001 (0.004)	-0.069*** (0.011)	0.076*** (0.016)
Latino × Share RM	0.017*** (0.004)	0.026*** (0.004)	0.014** (0.004)	0.005*** (0.001)	0.000 (0.002)	-0.007 (0.008)	0.037*** (0.01)
White × Share RM	-0.001 (0.005)	0.012* (0.005)	-0.020*** (0.005)	0.004* (0.002)	0.001 (0.002)	-0.001 (0.009)	0.018 (0.013)
N	2,511,150	2,511,150	2,511,150	2,511,150	2,511,150	2,511,150	1,573,787
Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Student-clustered standard errors shown in parentheses. * p<0.05, ** p<0.01, *** p<0.001. All columns include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. <i>GPA</i> , <i>Work Habits</i> , and <i>Cooperation</i> are standardized to the year level. <i>Held Back</i> and <i>Suspension</i> are dummy variables and indicate if a student has been held back or suspended for the given year. <i>ln(Days Absent)</i> is calculated using ln(x+1) to account for zeros. <i>Noncog. Index</i> calculated using exploratory factor analysis for students with math and ELA test scores, which restricts the number of observations. Standard errors for <i>Noncog. Index</i> are bootstrapped using 100 repetitions. <i>Share RM</i> is the share of race matched teachers a student has in a given year.							

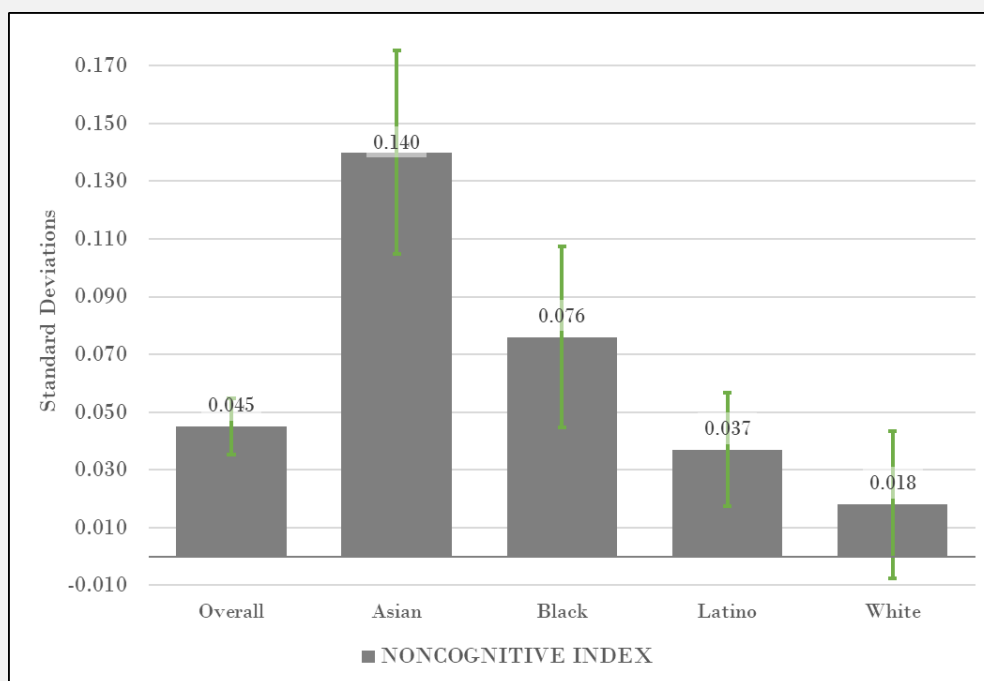
the effects found in Panel A for students of color. Race matched Asian, Black, and Latino students are expected to experience increases in each of these measures with Asian students expected to benefit the most. In fact, I find that, on average, Asian students benefit from race matching across each component with lower grade retentions, suspensions, and days absent. Black students are also expected to miss fewer school days when race matched. The overall findings of increased probabilities for race matched students to be held back seem to be driven by Latino and White students. In general, White students do not seem to benefit from having a larger share of race matched teachers. In fact, I find and expected race match effect of -0.020 sd in cooperation for White students. If White students do not benefit from race matching in terms of noncognitive skills, increasing the share of teachers of color in the teacher labor market will not detract from the development of these skills for White students.



## 1.4.2 Noncognitive Index

For context, Jackson (2018) estimates a *behavior index* using GPA, grade retention, absences, and suspensions. He finds that a standard deviation increase in the behavior index is expected to increase in the probability of graduating high school by 0.158. In a similar study, Petek and Pope (2021) find that a one standard deviation increase in the behavior index decreases the probability of dropping out by 0.020. These authors also generate a *learning skills index* using GPA, work habits, and cooperation and find a standard deviation increase to decrease the probability of dropping out by 0.040. Using these studies, I calculate back-of-the-envelope estimates that race matching has on the probability of dropping out through the expected changes in the noncognitive index.

**Figure 1.6 Race Match Effect on Learning Skills and Behavior**



Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12 with ELA and math test scores. *Noncognitive Index* is generated using an exploratory factor analysis with each noncognitive component, ELA, and math test scores. Standard errors are bootstrapped using 100 repetitions, and the whisker show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* shows the results for Equation 1.1 and was analyzed separately from the remaining groups. Equation 1.2 was used to analyze the race match effects by student race. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure corresponds to Column (7) of Table 1.3.

Figure 1.6 illustrates the overall and by student race results for the noncognitive index, which

captures student effort level, cooperation, and externalizing behaviors.<sup>13</sup> To account for the generated outcome variable, the standard errors are calculated using 100 bootstrap repetitions. The whiskers in the figure indicate a 95% confidence interval around each estimated race match effect. The first estimate shows the overall race match effect using Equation 1.1 and indicates an expected effect of 0.045 sd on the noncognitive index, which implies a decrease in the probability of dropping out between 0.0009 and 0.0018. Assuming a static, one-time effect of race matching and if all students were race matched at least once, these decreases represent between 225 to 500 students that dropped out of LAUSD during the period studied in this analysis.<sup>14</sup> The remaining coefficients are jointly estimated using Equation 1.2 and illustrate the impact that race matching has on the noncognitive index by student race. As with the individual noncognitive components, Asian students seem to benefit the most from race matching with an expected effect of 0.140 sd (a decrease in the probability of dropping out by 0.0028 to 0.0056). The noncognitive index for race matched Black and Latino students are also expected to increase by 0.076 sd and 0.037 sd, respectively. I do not find statistically significant effects for White students.

## 1.5 Course-Level Results for GPA, Work Habits, and Cooperation

The course-level analysis serves two purposes. First, the more granulated data allow for more precise estimates of the race match effect. In theory, if Equation 1.2 correctly identifies the race match effect, the coefficient on *Share RM* should be similar to the estimates of the effect of direct race matches on noncognitive outcomes. Second, since there is an additional level which students may race match (i.e., by course), this analysis allows for altering the level of fixed effects, which changes the necessary assumptions to identify the race match effect. This provides a test for the assumptions stated in Section 1.3.5 and is further explored in Section 1.6. A major limitation to using the course-level data is that several of the noncognitive outcomes (i.e., held back, suspensions, absences, and noncognitive index) are reported at the year level, which excludes them from this analysis.

---

<sup>13</sup>This figure corresponds to Column (7) of Table 1.3.

<sup>14</sup>Calculated using the 249,768 students that dropped out of LAUSD within the panel dataset from 2008 to 2018.

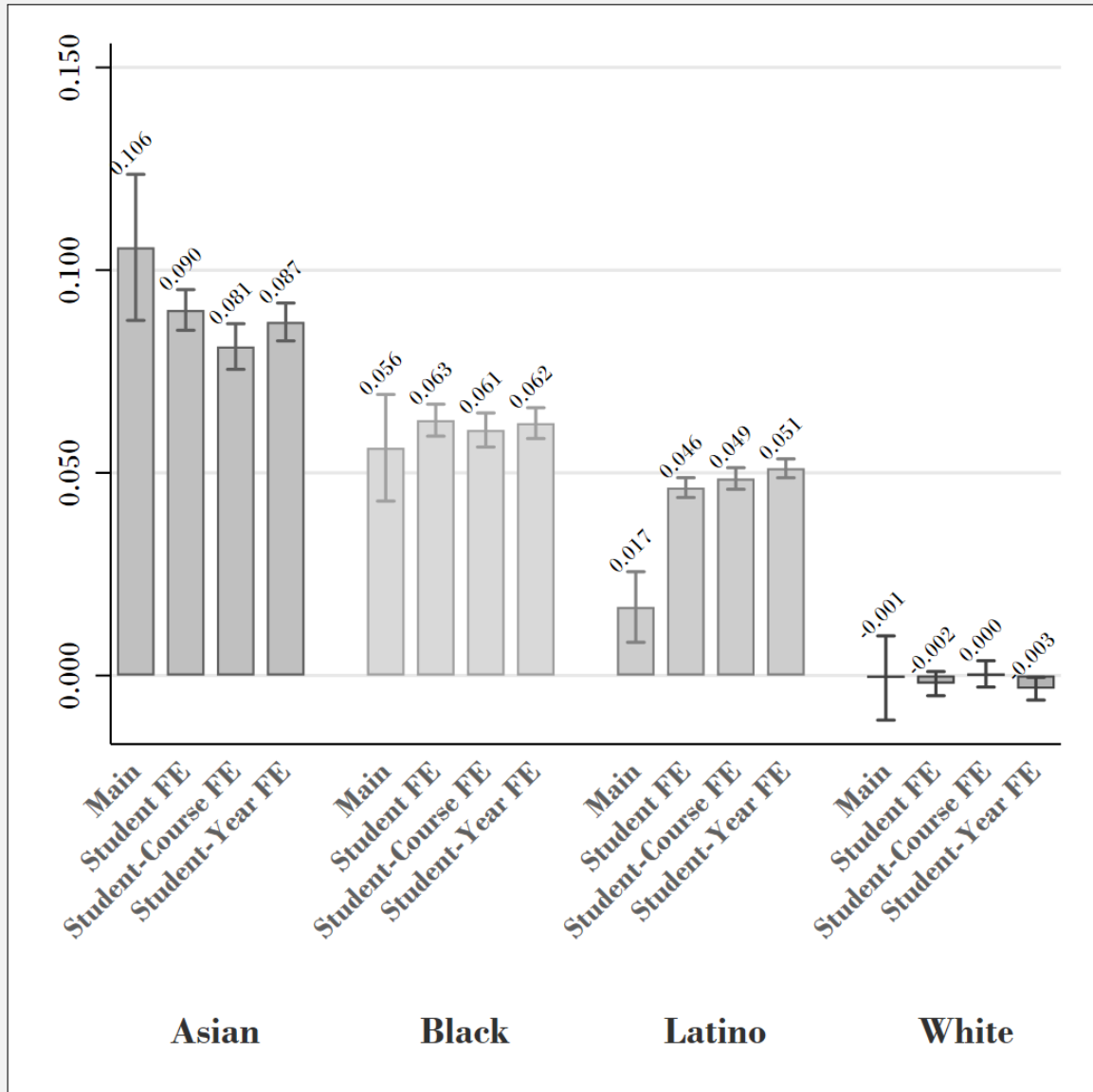
Throughout this study, I use the terms *course* and *subject* interchangeably. Table E.1 shows the percent of direct race matches by course/subject and student race. *History* also includes courses labeled as *social studies*, and *Health* includes *physical education* courses. The first row shows the overall race match percentages at the course-year level. The remaining rows of Table E.1 illustrate that certain races are more likely (compared to their overall race match percentage) to sort into subjects taught by a same race teacher. For example, on average, Asian students are directly race matched in math (21.6%) and science (18.8%) courses at a much higher rate than history (9.4%) or English language arts (10.5%) courses. Latino students match the race of their science teachers (25.7%) at a lower rate than their overall RM percentage (31.7%), and the same is true for White students and their math teachers (54.7% compared to 61.3%). Of course, it is important to take into account the racial distribution of teachers across courses as well. Table E.2 shows the percent of teachers across subjects by teacher race and reflects a similar pattern seen in the previous table. Asian teachers instruct math and science courses at a higher rate than they do for other courses. Similarly, Latino and White teachers teach science and math courses at a lower rate than they do for other courses.

Figure 1.7 illustrates the results for Equation 1.3 with Panel A showing the race match effects for GPA, Panel B for work habits, and Panel C for cooperation. For comparison, the main results (from Equation 1.2 and found in Table E.3) using the year-level data set are included. The second bar (identified by "Student FE") represents the coefficients and 95% confidence interval for Equation 1.3. The race match effects for each noncognitive component are similar to the main results, especially for Asian, Black, and White students. Though the effects in the main results and those found for the course-level analysis yield the same sign for Latino students, the course-level results are significantly larger. One possible explanation could be the differences in the each approach's ability to control for racial sorting bias across subjects. While both *ShareRM* and the direct race match use the course level data, *ShareRM* is an aggregation of a student's race matches across courses and within years. The inclusion of student fixed effects exploits the variation found within student and across years. While the approach to the course level analysis is similar, the student

fixed effects exploits the variation found in race matches across years *and* across courses, which mitigates the potential for racial sorting further if students are biased into choosing racial matched teachers only for particular subjects. However, it should be stated that, due to data limitations, the aggregated approach is the only viable option for outcomes reported at the yearly level (i.e., heldback, suspensions, days absent, and the noncognitive index).

**Figure 1.7 Direct Race Match Effect at Course Level**

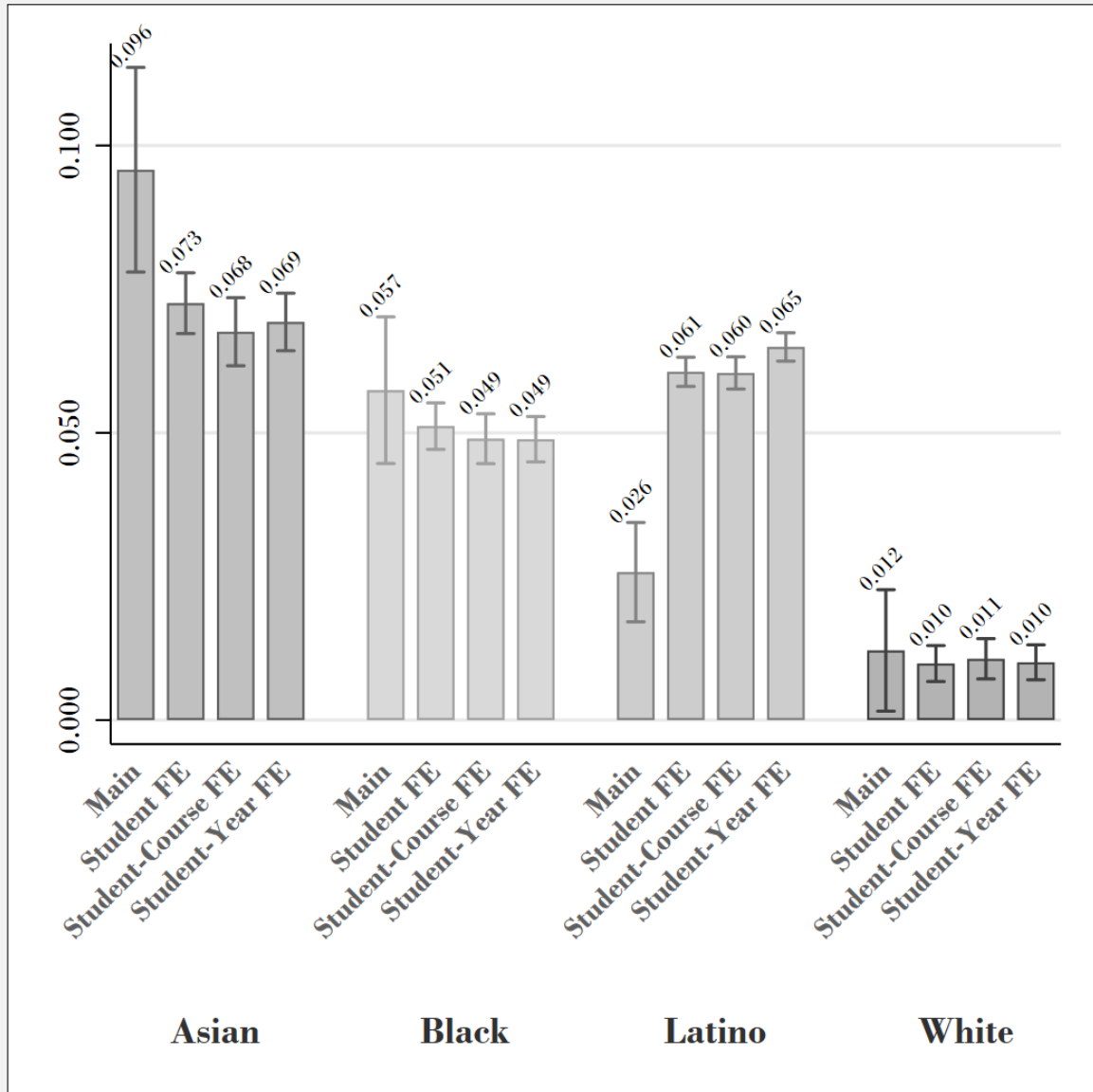
**Panel A: Course Level Analysis (GPA)**



Observations for the *Main* bars are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Observations for the *Student FE*, *Student-Course FE*, and *Student-Year FE* are at the student-year-course level for the same years and grades as the main sample. These coefficients were estimated using equation 1.2 with two changes. First, since these data are analyzed at the course-level, the race match effect captures a direct race match with the teacher. Second, the level of fixed effects were adjusted to match the given label. *GPA* (Panel A), *Work Habits* (Panel B), and *Cooperation* (Panel C) are standardized to the year level.

**Figure 1.7 (cont'd) Direct Race Match Effect at Course Level**

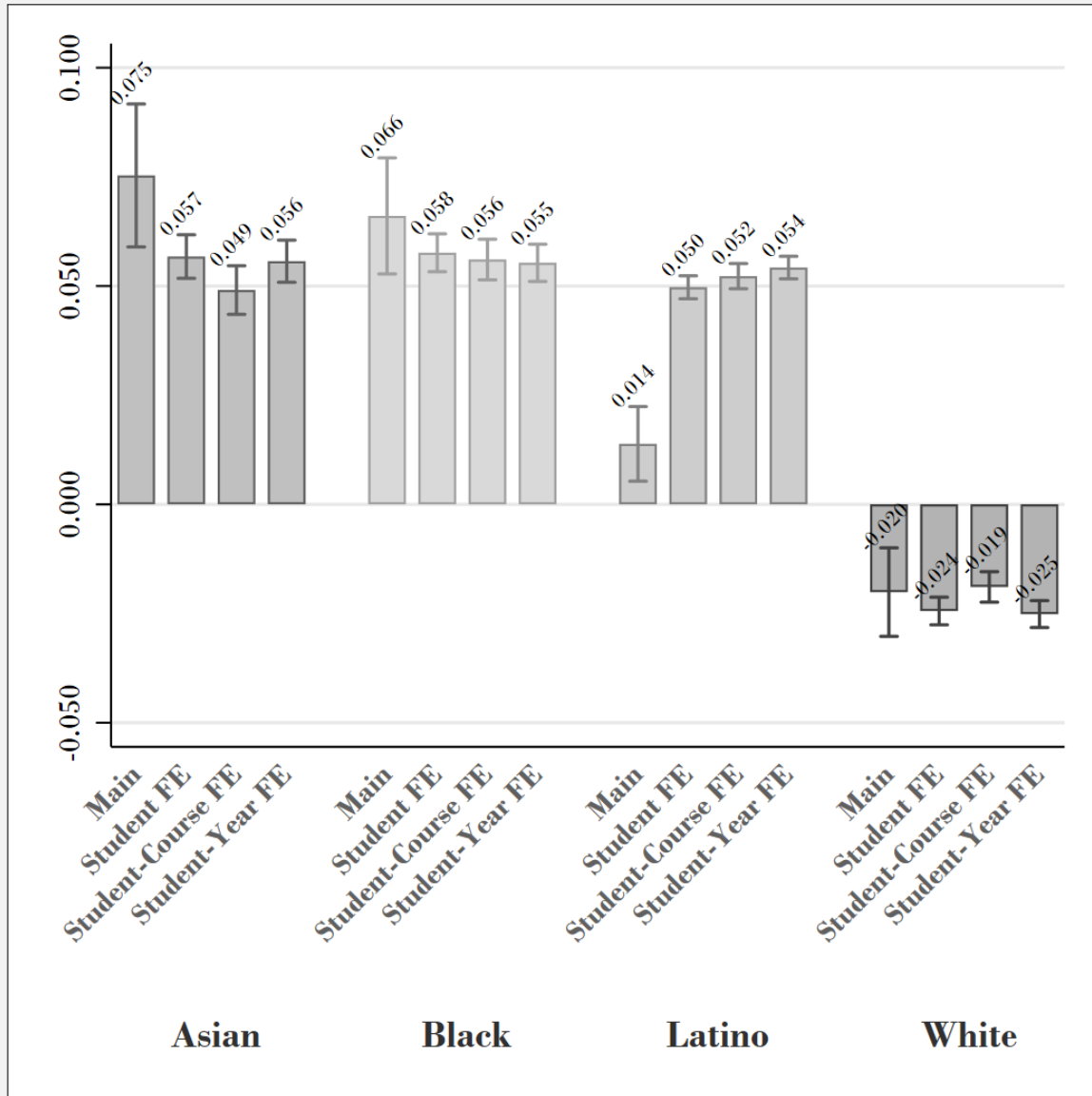
**Panel B: Course Level Analysis (Work Habits)**



Observations for the *Main* bars are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Observations for the *Student FE*, *Student-Course FE*, and *Student-Year FE* are at the student-year-course level for the same years and grades as the main sample. These coefficients were estimated using equation 1.2 with two changes. First, since these data are analyzed at the course-level, the race match effect captures a direct race match with the teacher. Second, the level of fixed effects were adjusted to match the given label. *GPA* (Panel A), *Work Habits* (Panel B), and *Cooperation* (Panel C) are standardized to the year level.

**Figure 1.7 (cont'd) Direct Race Match Effect at Course Level**

**Panel C: Course Level Analysis (Cooperation)**



Observations for the *Main* bars are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Observations for the *Student FE*, *Student-Course FE*, and *Student-Year FE* are at the student-year-course level for the same years and grades as the main sample. These coefficients were estimated using equation 1.2 with two changes. First, since these data are analyzed at the course-level, the race match effect captures a direct race match with the teacher. Second, the level of fixed effects were adjusted to match the given label. *GPA* (Panel A), *Work Habits* (Panel B), and *Cooperation* (Panel C) are standardized to the year level.

## 1.6 Robustness Checks

Students and teachers may sort by race for reasons that are correlated with the student's noncognitive abilities, which threatens the identification of the race match effect. I address this issue in Equation 1.1 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. I also include school-grade fixed effects, which mitigates the concern of students and teachers sorting across schools. However, a student's desire to sort in one year may not be the same as a following year. As further checks for sorting, I analyze three additional specifications, one for the course-level analysis using Equation 1.3 and two for the main analysis using Equation 1.1. First, I compare different levels of fixed effects by modifying Equation 1.3 for the course-level outcomes (i.e., GPA, work habits, and cooperation). Changing the level of fixed effects alters the necessary identifying assumptions and finding similar results between the different models mitigates this concern for sorting bias. Second, I include lagged outcomes in the regression model for Equation 1.1, which control for the case of students with similar levels of perceived noncognitive skills being sorted into the same classrooms. Finally, I provide a placebo test by replacing the *Share RM* variable in Equation 1.1 with the student's following year *Share RM*. I do not find evidence of sorting by race in any robustness check when taking into account student, teacher, and school characteristics, which provides evidence for Equations 1.1 and 1.3 being correctly specified.

### 1.6.1 Course-Level Altering 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 correctly



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 true 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-course for reasons that change every year, these methods would still be biased. As stated in the previous section, the noncognitive outcomes for the course-level analysis is restricted to GPA, work habits, and cooperation. In addition to Equation 1.3, the equations for this analysis are given by:

$$Y_{ict} = \beta_0 + \beta_{RM}(\mathbb{1}(Race_i) * RM_{ict}) + \mathbf{X}_{ict}\mathbf{\Gamma}_1 + \mathbf{Z}_{ict}\mathbf{\Gamma}_2 + \bar{\mathbf{X}}_{ict}\mathbf{\Gamma}_3 + \theta_{gs} + \theta_{ic} + \theta_t + \epsilon_{icgst}, \quad (1.4)$$

$$Y_{ict} = \beta_0 + \beta_{RM}(\mathbb{1}(Race_i) * RM_{ict}) + \mathbf{X}_{ict}\mathbf{\Gamma}_1 + \mathbf{Z}_{ict}\mathbf{\Gamma}_2 + \bar{\mathbf{X}}_{ict}\mathbf{\Gamma}_3 + \theta_{gs} + \theta_{it} + \theta_c + \epsilon_{icgst}, \quad (1.5)$$

where Equation 1.4 utilizes student-course ( $\theta_{ic}$ ) and year ( $\theta_t$ ) fixed effects to identify the race match effect by leveraging across year, within-course variation. Similarly, Equation 1.5 utilizes student-course ( $\theta_{it}$ ) and course ( $\theta_c$ ) fixed effects to identify the race match effect by leveraging within-year, across course variation. The standard errors for each model are clustered at the student-course and student-year level, respectively.

Figure 1.7 in the previous section compares the coefficients from these equations to those from the main analysis (using year-level data) and the course-level analysis from Equation 1.3 with Panel A showing the race match effects for GPA, Panel B for work habits, and Panel C for cooperation. Overall, comparing the results across each course-level model reveals that the race match effects are nearly identical. This provides evidence that the course-level results are robust to student-teacher sorting by race within schools and across subjects.

### 1.6.2 Course-Level Placebo Tests

An ideal placebo/spillover test for course-level race matching would be to analyze the race match effect by using their race match status for another course. However, students have multiple teachers and are enrolled in multiple courses/subjects, which makes it difficult to determine which other

course to use as the placebo. An alternative approach is to separately analyze the race match effect using each course as a placebo. I create a placebo sample which restricts the number of observations to students enrolled in each subject within a given year. Then, for all years in the panel and each of the five subjects, I create an indicator variable which marks the student as having a race match or not. Finally, I create two separate data sets by collapsing the placebo sample data by focusing on ELA and math outcomes. Using these data sets, I compare the race match effect on noncognitive outcomes in ELA and math courses to the placebo effects by using the race match status in the remaining courses. While creating the placebo sample, I compared any changes to the composition of student race-by-year, -by-grade, -by-subject, -by-race match status, and different combinations of each of these. The only noticeable difference between the two samples is that the grade distribution for the placebo sample is lower than the overall course-level sample due to older students not being enrolled in Health courses.

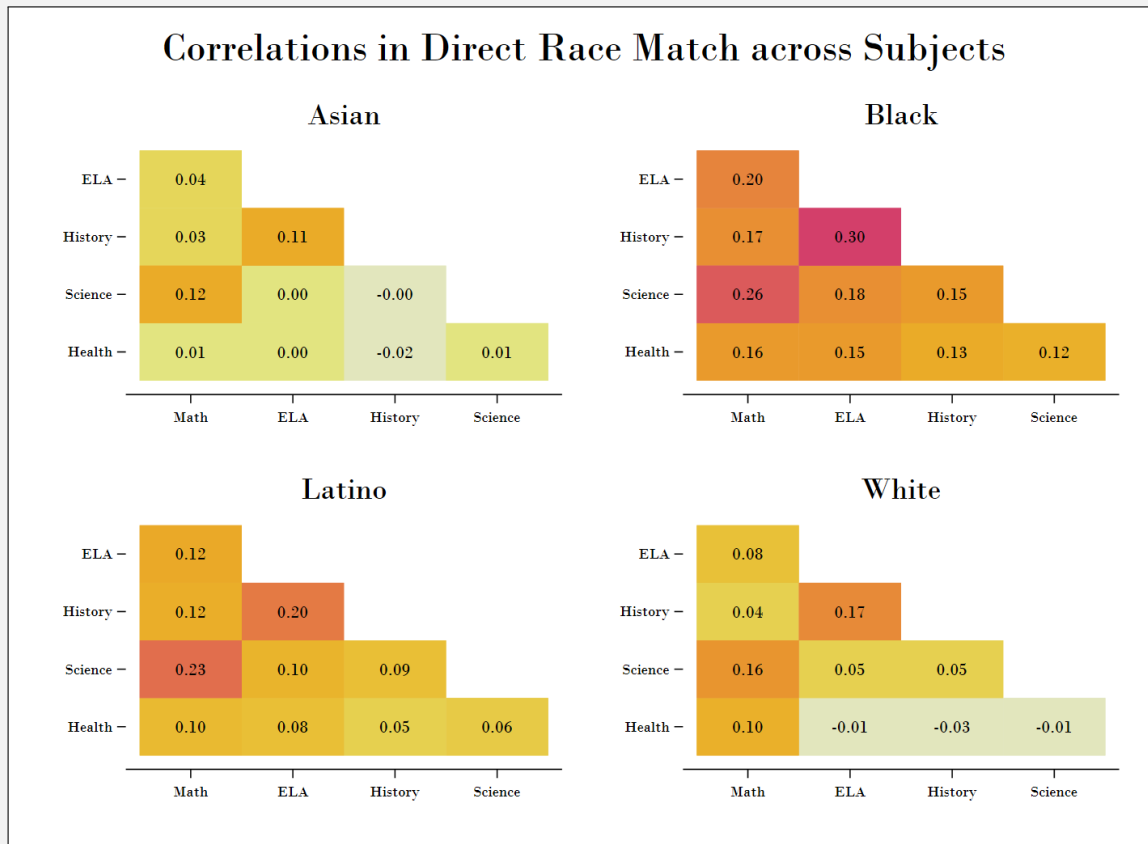
For context, Figure 1.8 illustrates the correlations in race match status by student race. For example, in the upper-left quadrant, the first cell shows a low correlation (0.04) between Math and ELA courses for Asian students. Across all races, math/science and ELA/history courses have the highest correlations in race match status. Similarly, the correlations for history/health pairings are relatively low for each race. Figure 1.9 illustrates the coefficients of the placebo test for ELA outcomes,<sup>15</sup> with Panel A showing the results for GPA, Panel B for Work Habits, and Panel C for Cooperation. In each of these figures, the non-placebo race match effect is statistically positive with the remaining estimates (i.e., the placebos) being statistically insignificant. One interesting pattern is, within student race, the placebo courses that yield significant positive results tend to have relatively low correlations in race matching, while the opposite is true for the significantly negative placebo effects. For example, the correlations in race matching for Asian students in ELA/math (0.04) and ELA/health (-0.02) are relatively lower than the correlations for math/science (0.12) and ELA/history (0.11). Focusing on ELA GPA (Figure 1.9 Panel A), the placebo race match effects for math and health are statistically positive, while the coefficient for history is negative

---

<sup>15</sup>See Figure E.1 for the course-level placebo test for math outcomes.

(though statistically insignificant). While more analysis would need to occur to formally establish this relationship, students may benefit from a positive spillover effect in noncognitive outcomes towards subjects where they tend to not be race matched.

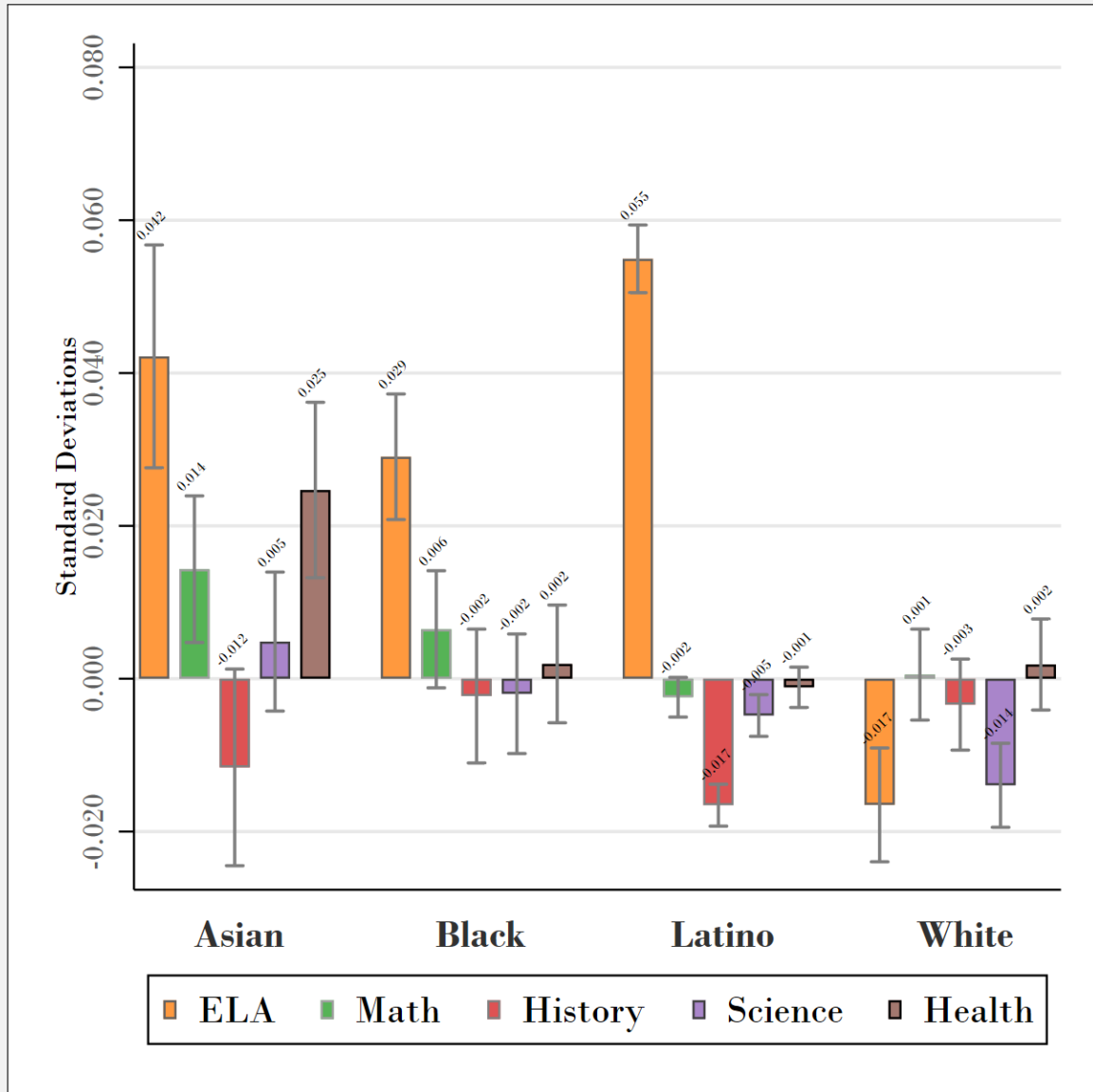
**Figure 1.8** Correlation of Direct Race Matches across Courses by Student Race



Observations are at the student-year-course level. The Main rows represent the percent of each race across grades from students in the main sample. The Placebo rows represent the percent of each race across grades from students in the placebo sample. The difference between the two sample percentages is presented below the black lines with red values indicating a higher percentage in the grade-race combination for the placebo sample. The placebo sample consists of all student-year observations from the main sample enrolled in English, Math, History, Science, and Health/P.E. courses. The final column represents the total number of student-years across grades for each sample. N=16,921,981 for the main sample. N= 10,324,132 for the placebo sample.

**Figure 1.9 Race Match Effects by Placebo Race Match Status (ELA)**

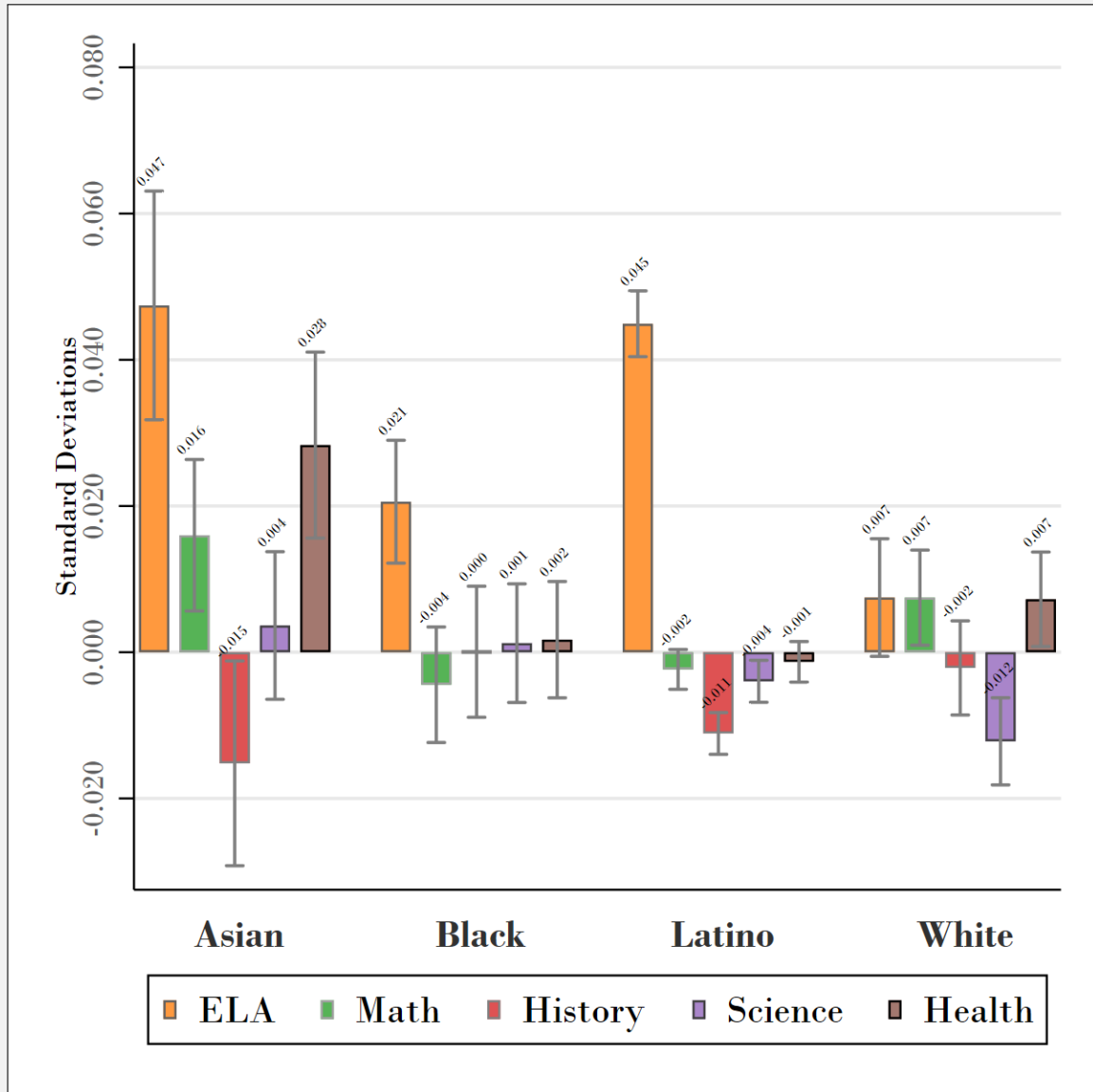
**Panel A: Course Level Placebo Analysis (GPA)**



Placebo sample consists of student-year observations enrolled in each English, Math, History, Science, and Health/P.E. courses within the same year. Observations are restricted from the placebo sample to the subject listed in the title. Each bar represents the coefficient from 1.2 using the race match indicator for the given subject. The titled subject is the estimated race match effect, and the remaining estimates are placebo tests. The whiskers show a 95% confidence interval around each estimate. All regressions include school-grade, student, subject, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the grade-subject-year level. Panel A shows the results for GPA; Panel B for Work Habits; and Panel C for Cooperation.

**Figure 1.9 (cont'd) Race Match Effects by Placebo Race Match Status (ELA)**

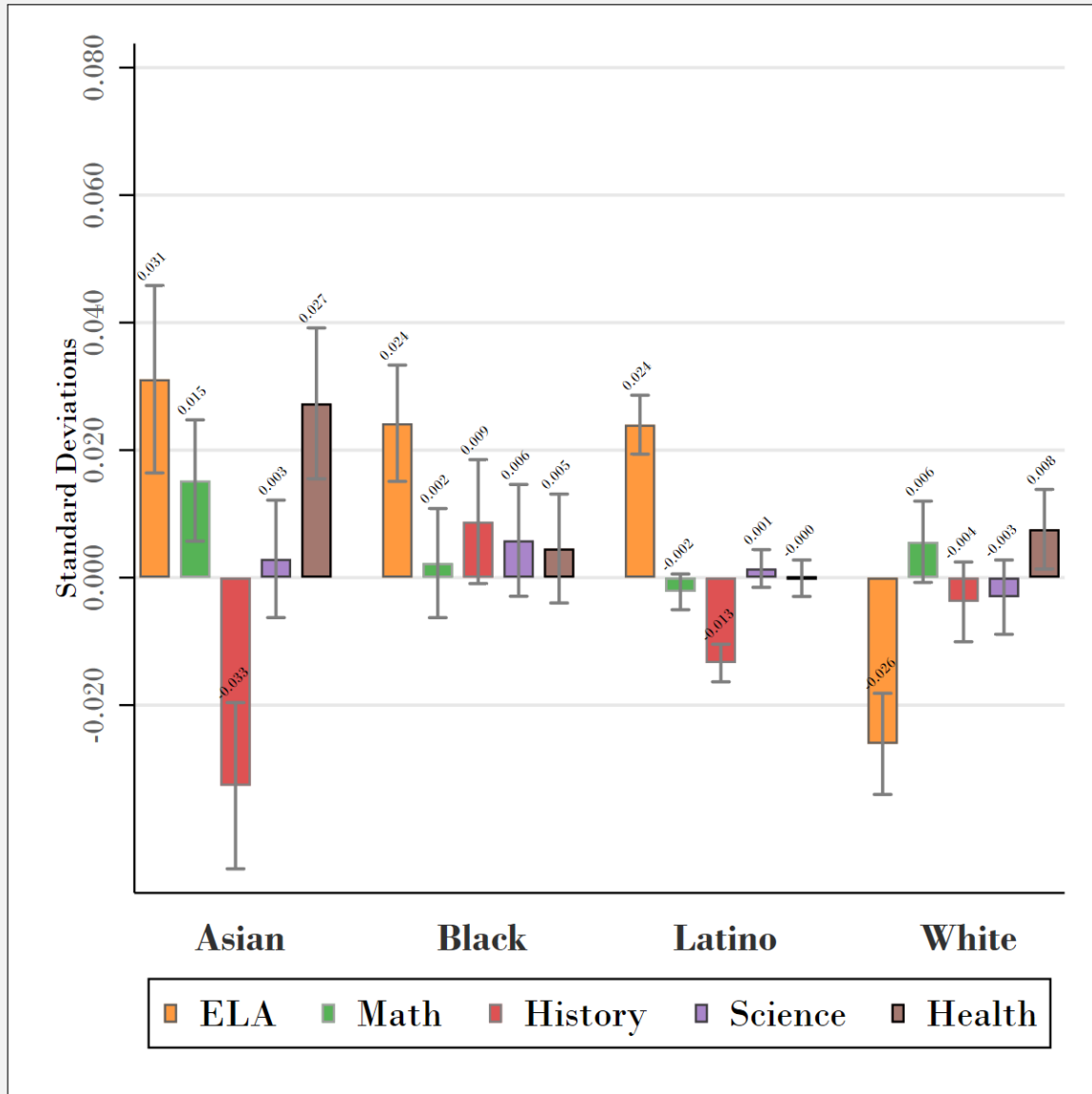
**Panel B: Course Level Placebo Analysis (Work Habits)**



Placebo sample consists of student-year observations enrolled in each English, Math, History, Science, and Health/P.E. courses within the same year. Observations are restricted from the placebo sample to the subject listed in the title. Each bar represents the coefficient from 1.2 using the race match indicator for the given subject. The titled subject is the estimated race match effect, and the remaining estimates are placebo tests. The whiskers show a 95% confidence interval around each estimate. All regressions include school-grade, student, subject, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the grade-subject-year level. Panel A shows the results for GPA; Panel B for Work Habits; and Panel C for Cooperation.

**Figure 1.9 (cont'd) Race Match Effects by Placebo Race Match Status (ELA)**

**Panel C: Course Level Placebo Analysis (Cooperation)**



Placebo sample consists of student-year observations enrolled in each English, Math, History, Science, and Health/P.E. courses within the same year. Observations are restricted from the placebo sample to the subject listed in the title. Each bar represents the coefficient from 1.2 using the race match indicator for the given subject. The titled subject is the estimated race match effect, and the remaining estimates are placebo tests. The whiskers show a 95% confidence interval around each estimate. All regressions include school-grade, student, subject, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the grade-subject-year level. Panel A shows the results for GPA; Panel B for Work Habits; and Panel C for Cooperation.

### 1.6.3 Lagged Outcomes

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. (2014b) 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. For each noncognitive outcome, I include a lagged measure of the outcome to Equation 1.2 resulting in the following equation:

$$Y_{it} = \beta_0 + \beta_1 Y_{it-1} + \beta_{RM}(ShareRM_{it}) + \mathbf{X}_{it}\Gamma_1 + \bar{\mathbf{Z}}_{it}\Gamma_2 + \bar{\mathbf{X}}_{it}\Gamma_3 + \theta_i + \theta_{gs} + \theta_t + \epsilon_{igst}. \quad (1.6)$$

Table 1.4 compare the race match effects between the main specification and equation 1.6 for each noncognitive outcome.<sup>16</sup>

**Table 1.4 Race Match Effects including Lagged Outcome**

Panel A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Overall	GPA	Work Habit	Cooperation	Held Back	Suspension	ln(Days Absent)	Noncog. Index
Share RM	0.025*** (0.003)	0.033*** (0.003)	0.020*** (0.003)	0.002 (0.001)	-0.001 (0.001)	-0.021*** (0.005)	0.054*** (0.007)
Panel B	(1)	(2)	(3)	(4)	(5)	(6)	(7)
By Race	GPA	Work Habit	Cooperation	Held Back	Suspension	ln(Days Absent)	Noncog. Index
Asian × Share RM	0.106*** (0.011)	0.100*** (0.011)	0.085*** (0.010)	-0.009** (0.003)	-0.003 (0.003)	-0.114*** (0.022)	0.144*** (0.024)
Black × Share RM	0.064*** (0.009)	0.071*** (0.009)	0.064*** (0.009)	0.000 (0.003)	0.001 (0.005)	-0.071*** (0.015)	0.141*** (0.020)
Latino × Share RM	0.011* (0.006)	0.023*** (0.006)	0.016** (0.005)	0.003 (0.002)	-0.003 (0.002)	-0.008 (0.010)	0.027* (0.013)
White × Share RM	0.009 (0.007)	0.013 (0.007)	-0.017* (0.007)	0.005* (0.002)	0.002 (0.003)	0.005 (0.012)	0.022 (0.017)
N	1,717,706	1,717,706	1,717,706	1,717,706	1,717,706	1,717,706	767,291

Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered. \* p<0.05, \*\* p<0.01, \*\*\* p<0.001. Coefficients based on Equation 1.1 including the a lagged outcome into the regression model. All models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the year level. *Held Back* and *Suspension* are dummy variables and indicate if a student has been held back or suspended for the given year. *ln(Days Absent)* is calculated using ln(x+1) to account for zeros. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year.

Though the results in this table are similar to those found in Table 1.3, a few differences stand out. While the overall effect size (Panel A) effect for held back is the same as before (0.02 percentage points), the race match effect when controlling for a student's lagged outcome is statistically indistinguishable from zero, though the original effect size is fairly small. Also, the

<sup>16</sup>Figure F.1 is the figure analogue for Panel A in Table 1.4 and is found in Appendix F.

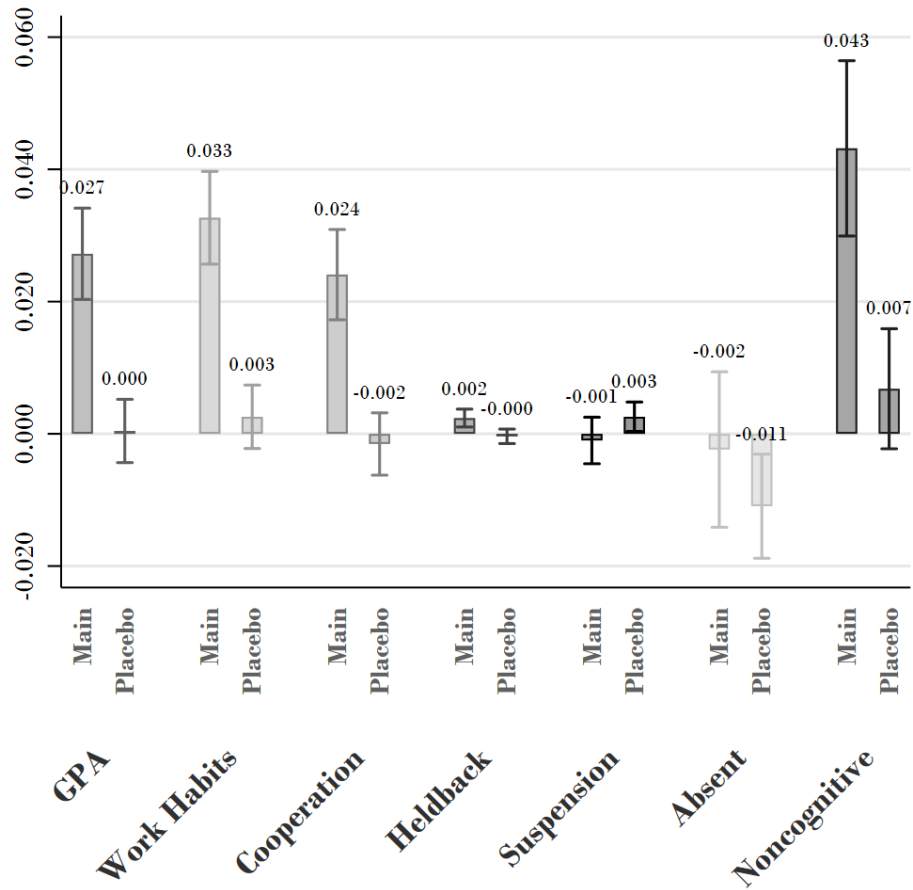
Black race match effect for GPA (0.056 sd compared to 0.0064 sd), work habits (0.057 sd and 0.071 sd), and the noncognitive index (0.076sd and 0.141 sd) are larger once controlling for a student's previous outcome. In general, the similarities between the estimated race match effects across both model specification evidences that Equation 1.1 correctly controls for student-teacher racial sorting within grades and schools.

#### **1.6.4 Placebo Test using Following Year Race Match**

In theory, a student's outcome should not be impacted by their race match status in the following year. Figure 1.10 illustrates a placebo test for the overall race match effect by comparing the results for Equation 1.1 to those when replacing the share race match variable with a student's share of race matched teachers in the following year,  $ShareRM_{igst+1}$ . To ensure a proper comparison, the sample is restricted to observations available in both regressions. Each coefficient is interpreted as before (i.e. change in standard deviations for GPA, Work Habits, and Cooperation, change in the probability of receiving a Suspension for Heldback, or the change in the percent of days Absent). The estimated placebo effects for GPA, Work Habits, Cooperation, Heldback, and Noncognitive are statically indistinguishable from zero, which provides some evidence for the validity of the main results found for these components. For Suspension and Absent, the respective race match effects and placebo effects are not statistically different from each other. However, when the sample is restricted to observations available for both regression models, the non-placebo effect sizes for both components are statistically insignificant. Notably, for Absent, the placebo result is statistically negative, but only by a relatively small amount (1.1%).



**Figure 1.10** Placebo Test using following Year Share Race Match



The first bars represent the coefficients for  $ShareRM_{igst}$  from equation 1.1. The second bars estimate a placebo test by replacing  $ShareRM_{igst}$  with the following year's share race match,  $ShareRM_{igst+1}$ . The same sample was used in both regressions with  $N = 1,087,824$  for all outcomes except for Noncognitive, which had  $N = 813,322$  due to the sample restriction for students with ELA and math test scores. The coefficients for GPA, Work Habits, Cooperation, and Noncognitive are interpreted as a change in standard deviations. The coefficients for Held Back and Suspension are interpreted as a change in percentage points. The coefficient on  $\ln(\text{Absent})$  is interpreted as a percentage change.  $\ln(\text{Days Absent}+1)$  was used to account for zero days absent. The whiskers represent a 95% confidence interval around the estimated effect. Noncognitive uses bootstrapped standard errors with 100 repetitions.

## 1.7 Non-linear Race Match Effect

The race match effect may differ based on a student's current exposure to similarly raced teachers. The impact of race matching may be larger for students with lower shares of race matched teachers and diminish as this share increases. Figures 1.11 and 1.12 illustrate the non-linear effects of

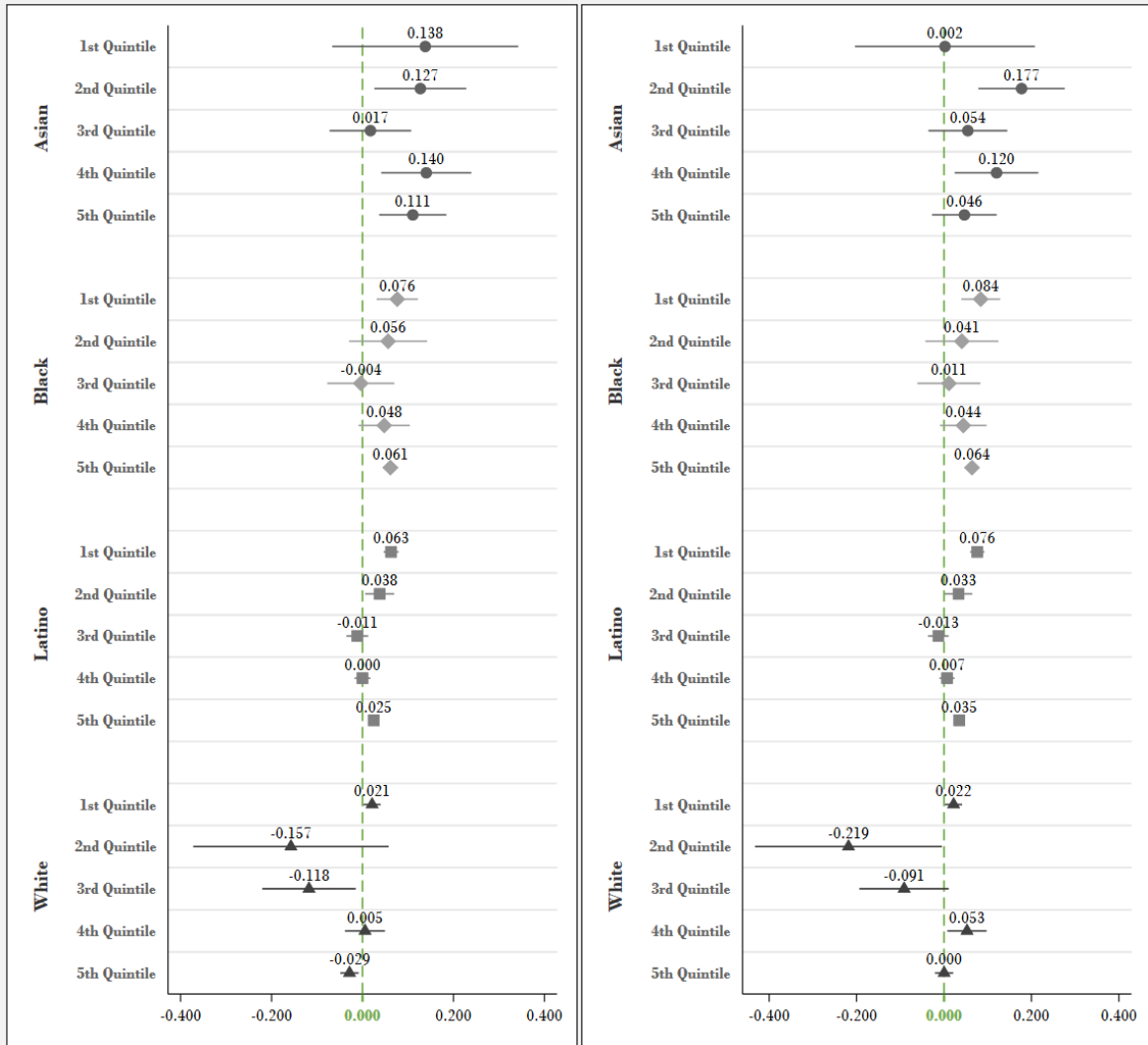
race matching on noncognitive outcomes by separating the observations into quintiles of the *Share RM* variable by year. By student race, each point represents the estimated race match effect for students in the given quintile with *1st Quintile* indicating students with the lowest number of race matched teachers. The whiskers represent a 95% confidence interval for each estimate. Given the similarities between the results for bootstrapped and non-bootstrapped standard errors and the computational requirements, I do not use bootstrapped standard errors for the noncognitive index analysis.

Panels A, B, and C of Figure 1.11 and Figure 1.12 show a similar pattern within each student race for GPA, work habits, cooperation, and the noncognitive index respectively. For the first quintile (i.e., lowest share of race matched teachers), the race match effect is significantly positive for Black and Latino students, but this effect decreases as the share of race matched teachers increases up to the third quintile. Then, the race match effect increases again. One possible explanation for this pattern is that Black and Latino students with lower shares of race matched teachers benefit through the mechanisms explained in Section 1.2. However, these effects seem to diminish as the student's exposure to race matched teachers increases. That is, until the student is race matched to a large enough share of their teachers that the race match effect begins to increase again. This may indicate that student's with higher shares of race matched teachers benefit from *homophily*. That is, students with higher shares of race matched teachers may see inflated scores on GPA, work habits, and cooperation due to teacher bias. Again, in this case, the race match effect is no longer interpreted as having a direct impact on a student's noncognitive ability, but may still indirectly affect the student through positive affirmation. Generally speaking, there do not seem to be any non-linear effects for the remaining outcomes. Two exceptions are for suspensions and absences in the lower quintiles for Asian and Black students. In each of these, race matched students seem to be less likely to be suspended or be absent, but the effect diminishes as the share of race match teachers increases.

**Figure 1.11 Race Match Effects by Quintile of Share RM**

**Panel A: GPA (standard deviations)**

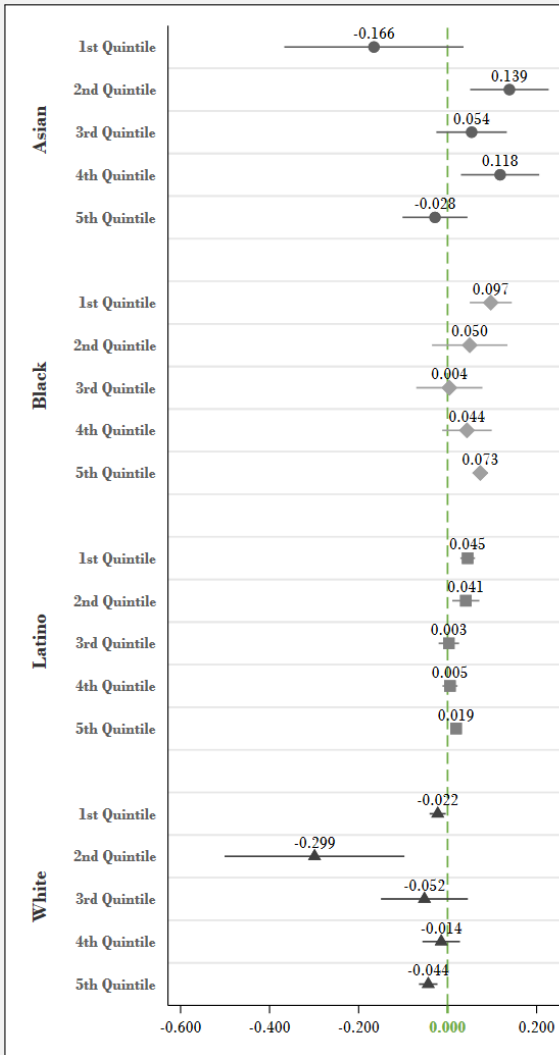
**Panel B: Work Habits (standard deviations)**



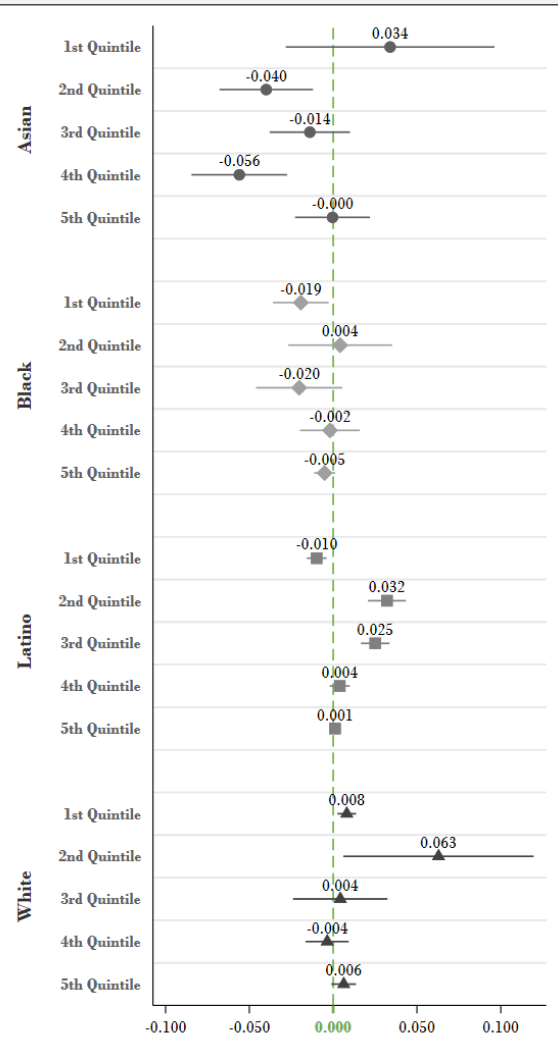
Students are binned into each quintile based on the distribution of share race match. The first quintile represents students with no to low shares of race matched teachers, and the fifth quintile represents students with higher shares of race matched teachers. Each point is illustrated with a 95% confidence interval and represents the share race match effect with the dotted line indicating no effect. The groups by student race are estimated using equation 1.2. The following lists each panel and outcome pairing. Panel A: GPA; Panel B: Work Habits; Panel C: Cooperation; Panel D: Held Back; Panel E: Suspension; Panel F: ln(Days Absent).

**Figure 1.11 (cont'd) Race Match Effects by Quintile of Share RM**

**Panel C: Cooperation (standard deviations)**



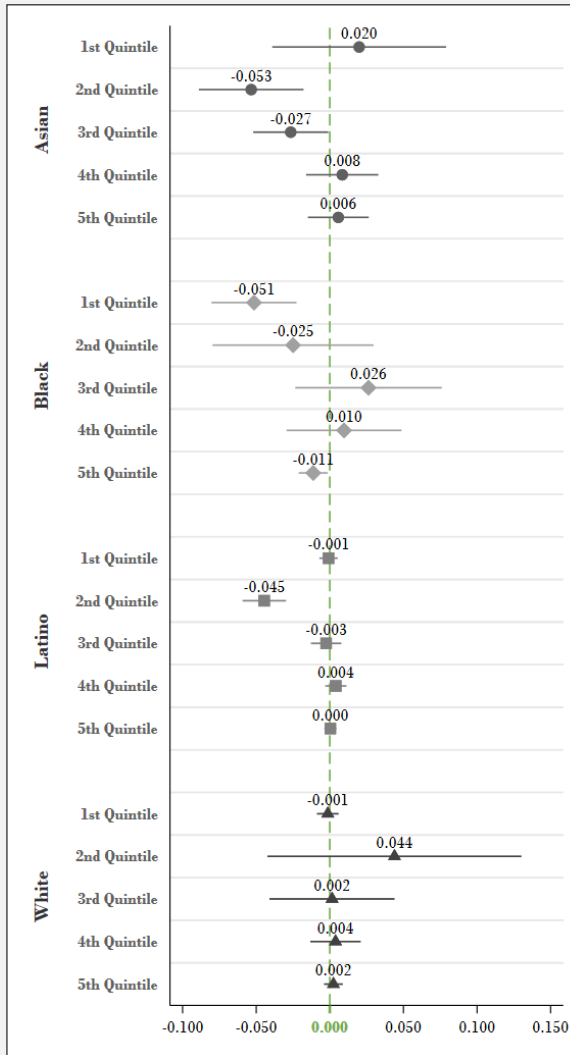
**Panel D: Held Back (probability)**



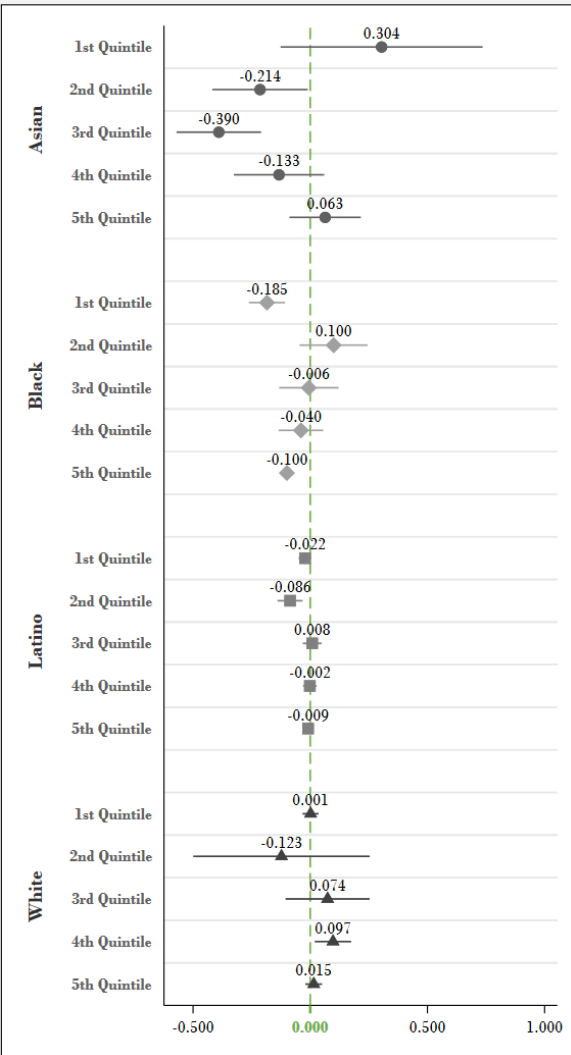
Students are binned into each quintile based on the distribution of share race match. The first quintile represents students with no to low shares of race matched teachers, and the fifth quintile represents students with higher shares of race matched teachers. Each point is illustrated with a 95% confidence interval and represents the share race match effect with the dotted line indicating no effect. The groups by student race are estimated using equation 1.2. The following lists each panel and outcome pairing. Panel A: GPA; Panel B: Work Habits; Panel C: Cooperation; Panel D: Held Back; Panel E: Suspension; Panel F: ln(Days Absent).

**Figure 1.11 (cont'd) Race Match Effects by Quintile of Share RM**

**Panel F: Suspension (probability)**

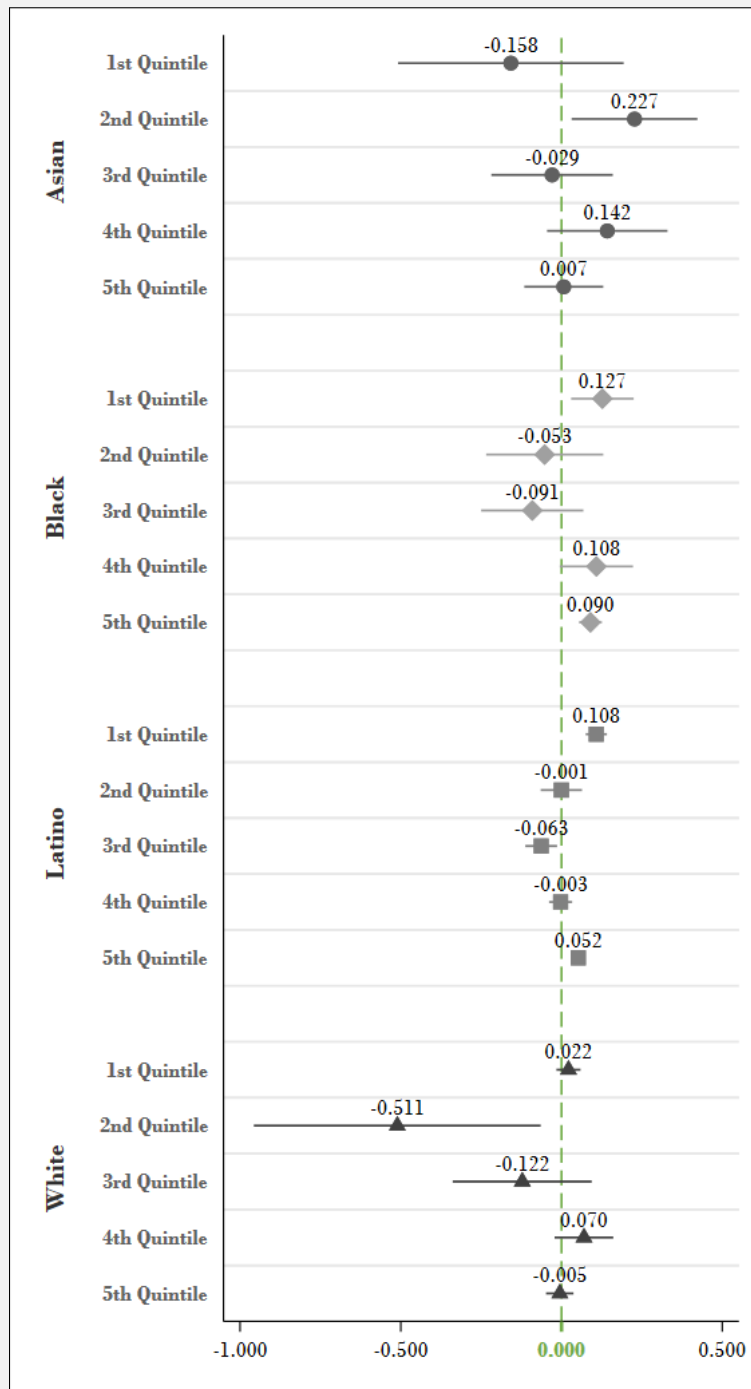


**Panel G: ln(Days Absent)**



Students are binned into each quintile based on the distribution of share race match. The first quintile represents students with no to low shares of race matched teachers, and the fifth quintile represents students with higher shares of race matched teachers. Each point is illustrated with a 95% confidence interval and represents the share race match effect with the dotted line indicating no effect. The groups by student race are estimated using equation 1.2. The following lists each panel and outcome pairing. Panel A: GPA; Panel B: Work Habits; Panel C: Cooperation; Panel D: Held Back; Panel E: Suspension; Panel F: ln(Days Absent).

**Figure 1.12 Race Match Effects on Noncognitive Index by Quintile of Share RM**



Students are binned into each quintile based on the distribution of share race match. The first quintile represents students with no to low shares of race matched teachers, and the fifth quintile represents students with higher shares of race matched teachers. Each point is illustrated with a 95% confidence interval and represents the share race match effect with the dotted line indicating no effect. The groups by student race are estimated using equation 1.2. The following lists each panel and outcome pairing. The noncognitive index was generated using an exploratory factor analysis with the components found in Figure 1.11, ELA, and math test scores.

## 1.8 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.

When students are placed with a teacher of the same race, they may be less likely to worry about exhibiting negative racial stereotypes. I find that students of color in grades 6 through 12 benefit from increased race matches with improvements to noncognitive outcomes such as GPA, work habits, and cooperation. I also find fewer missed days of school for race matched Asian and Black students. These impacts are robust to several model specification and placebo tests. I also find non-linear race match effects indicating that students of color with lower shares of race matched teachers benefit the most from increased exposure to teachers of the same race.

Significant policy implications of this study are to reinforce the need for school districts to measure and collect noncognitive outcome data for students and to further diversify the teacher workforce. Improvements to noncognitive outcomes is associated with higher probabilities of graduating high school. It is important to further study influences on noncognitive outcomes, which can only occur if the data is available for research. Secondly, while the racial composition of students continues to diversify, White teachers remain the vast majority. In fact, over 80 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 be achieved by implementing and expanding policies to attract and train more people of color into teaching.

## **CHAPTER 2**

### **THE EFFECT OF SAME RACE TEACHERS AND FACULTY ON STUDENT TEST SCORES**

#### **2.1 Disclaimer**

This chapter was co-authored with Ijun Lai (ILai@mathematica-mpr.com), who has approved that this work be included as a chapter in my dissertation.

#### **2.2 Introduction**

Students benefit from being placed with a teacher of the same race across many outcomes, such as improved test scores (Dee, 2005; Egalite et al., 2015; Joshi et al., 2018), high school graduation rates (Gershenson et al., 2018), disciplinary outcomes (Lindsay and Hart, 2017; Shirrell et al., 2021), learning skills (Wood, 2022), as well as access to resources such as placement in gifted classes (Grissom et al., 2015). However, while the student population rapidly continues to diversify, the diversification of the teaching corps continues to lag behind. For example, the proportion of Latino student enrollment in public schools has increased from 11 to 27 percent in just the last two decades. In contrast, share of Latino public school teachers during this same period has increased from 3 to only 9 percent (Pew Research Center, 2021). If the disparity between student and teacher racial distributions continues to grow, students of color may find it more difficult to benefit from direct student-teacher race matching. However, it may still be possible for students to benefit from same-race teachers even if they are not placed in the same classroom. A growing number of studies highlight the importance of teacher peer effects (e.g., Jackson and Bruegmann, 2009; Sun et al., 2017; Elstad et al., 2020). These studies document the professional development and growth of teachers as they are exposed to higher and lower quality peers, proving that teachers can and do actively learn from one another. Correspondingly, this literature raises the possibility that increasing teacher diversity can influence how incumbent teachers interact with and teach students



of color. In other words, hiring an additional teacher of color could have a larger impact on students of color that is not being measured in the existing race match literature. For the sake of clarity, from a student's perspective, we refer to teachers with whom the student does not share a class as *faculty*. Students may benefit from faculty of the same race in multiple ways. First, teachers may learn from each other on how to improve instruction towards students from different backgrounds. Also, as documented from existing race match literature, students may benefit from the *role model effect* if students interact with same-race faculty members in meaningful ways. An increase in the share of faculty members of color may also signal a change in the school's culture, as facilitated by the principal, to improve educational outcomes for students of color. More detail on each of these potential mechanisms is provided in the subsequent section of this paper.

In this study, we bring together the race match literature and teacher peer effect literature to gain a more comprehensive understanding of how increasing the share of teachers of color can benefit schools beyond a direct race match with students. We utilize a rich set of administrative data from the Los Angeles Unified School District (LAUSD) to see what effect, if any, faculty-student race matching has on a student's test scores. We also explore how this effect differs across student race and how it compares to the more commonly measured direct race match effect.

LAUSD is uniquely situated to answer our research questions. First, the majority of race match papers either focus on nationally representative samples, which tend to have few students of color, or use administrative data from southern states (i.e. Tennessee, North Carolina, Florida). It is unclear whether similar findings would materialize in urban school districts, which have a greater share of minority students. Second, Los Angeles is one of few urban school districts that have significantly increased the proportion of teachers of color, particularly Latino teachers. For this reason, we also contribute to the direct race match literature by studying the teacher-student race match effect for Latino students and compare the results with those found for Black students.

We find that the average students' math scores improve when both their math teacher and other faculty members at their grade level share their race. Increases in same race faculty alone is associated with lower math scores. Conversely, we find that same-race faculty improve students'

English language arts (ELA) test scores, and that this effect is even larger when both the ELA teacher and an increased number of faculty are the same race as the student. When broken down by racial group, it appears these results are largely driven by the Latino student population. Latino students benefit in both math and ELA when both their teachers and a greater share of faculty are also Latino. Other racial groups (i.e., Asian, Black and White) see more mixed results.

## **2.3 Theoretical Mechanisms and Relevant Literature**

### **2.3.1 Student-teacher race match effect on test scores**

There are now a handful of rigorous papers showing the impact of student-teacher race match on student outcomes—particularly for Black students. The most influential paper comes from Dee (2004)’s analysis of Project STAR in Tennessee. This project randomly assigned kindergarten through third grade students to teachers, thereby allowing researchers to analyze the causal impact of race match on student test scores. Dee found a 3 to 6 percentage point increase in test scores for race matched Black students. Since this paper, several studies have used different empirical strategies (predominantly various fixed effects models) and generally found similar results for race matched Black students across different contexts (i.e., grades, states, outcomes) (e.g., Egalite et al., 2015; Joshi et al., 2018). One potential mechanism for how a direct student-teacher race match may improve test scores is that being assigned to a teacher of the same race could reduce stereotype threat. When students fear behaving in ways that confirm negative stereotypes, they may underperform in academic settings due to an increase in anxiety (e.g., Steele and Aronson, 1997). When students are placed with a teacher of the same race, students of color may feel more comfortable participating in class or asking for help, which could improve their test scores.

Same race estimates for Latino students is much more limited, and results are mixed. Buddin and Zamarro (2009b) use data from Los Angeles and find no evidence of any race match benefits for Latino students in elementary school, but some evidence of gains for Latino students in middle school math classes. Using nationally representative data, Jennings and DiPrete (2010) find impacts in math for Latino students who were race matched in kindergarten through third grade. Similar to

Buddin & Zamarro, their results suggest no race match benefits for elementary students in reading scores. On the other hand, Egalite et al. (2015) utilizes administrative data from Florida and finds significantly negative impacts of race match for Latino students. This finding is consistent across both math and reading, as well as across grade level. However, it is unclear whether these results generalize beyond Florida. Florida's Latino population is very diverse and composed of Cuban, Puerto Rican, Mexican, Dominican and other Latino sub-populations. While data limitations may indicate that a Latino teacher with Cuban origin is race matched to a Latino student with Mexican origin, differences in the cultural identities of these two individuals may change the efficacy of a perceived, but perhaps not felt, "race match." Our study is based in California, where 84% of the Latino population reports their origin as Mexican (Pew Research Center, 2014). Consequently, we believe our study is uniquely situated to shed light on how race matches influence students of Mexican origin.

To the best of our knowledge, there are only two papers that even marginally address race match in Los Angeles (Buddin and Zamarro, 2009b,a). Both papers use student and teacher fixed effects to examine how teacher qualifications are related to student achievement, but also briefly examine the possibility of race match benefits. Overall, the authors do not find strong evidence of race match differences. There is some evidence of math gains for race matched Black students in elementary grades and, as previously mentioned, race matched Latino middle school students. However, these papers are limited in at least three important ways. First, these papers focus on data from 2000-2004 and 2000-2007. This study period misses the increase in Latino new teacher hires that seemingly began in 2009. Paired with the lower Latino teacher exit rates, the overall share of teachers who are Latino increased by seven percentage points between 2002 and 2011 (Institute, 2015). Second, this set of papers do not control for school characteristics. This omission may bias estimates since the literature has shown that teachers of color are more likely to teach at higher poverty and lower achieving schools with higher levels of teacher turnover (e.g., Achinstein et al., 2010; Hanushek, 2004). Finally, the authors do not address the role English Language Learner status may play in estimating race match effects for Latino students.

### **2.3.2 Teacher peer effects, school culture, and other-teacher/staff**

While much of the previous literature focuses on a direct student-teacher race matching, it ignores the influence teachers of color may have on other teachers and on students that may not be assigned to their own classrooms. The idea that teachers may learn from each other has been the focus of a growing literature which documents the importance of teacher peer effects. Specifically, these studies highlight teacher collaboration, where teachers are learning from one another to develop better teaching practices. These collaborations can occur formally, via professional development sessions or team meetings, or informally, such as in the teachers' lounges over lunch (e.g., Mawhinney, 2010; Spillane, 2012; Sun et al., 2013; Frank et al., 2014). Empirical studies find that the introduction of more effective teachers to a school can increase the performance of the other teachers in grade-level teams (Jackson and Bruegmann, 2009; Sun et al., 2017).

Beyond educational content and teaching materials, teachers may work with one another to learn how to better connect with their students and teach them more effectively. Part of this knowledge gap may stem from teachers being demographically different from their students. In these scenarios, it may be particularly beneficial for teachers to consult and learn from teachers of color (e.g., Ladson-Billings, 2004; Lim, 2006; Welsh and Little, 2018). As more teachers of color join a school, incumbent teachers of color may find it easier to informally build support and resources for students of color (Milner, 2006).

Higher proportions of teachers of color may also result in institutional changes or policy shifts that create more productive learning environments for students of color. For example, studies suggest that having a higher proportion of Black teachers in a school is associated with lower rates of exclusionary discipline practices (i.e., suspensions and expulsions) and higher rates of alternative disciplinary practices that are focused on student growth and learning (Grissom et al., 2009; Meier and Stewart Jr, 1992; Rocha and Hawes, 2009; Roch et al., 2010). These policy shifts can result in concrete resources for students. Since students often need teacher referrals to access extra resources, such as gifted classes or needed special education services, teachers are important gatekeepers. Schools with higher proportions of teachers and staff of color are associated with

increased students of color in these programs (e.g., Nicholson-Crotty et al., 2011). For example, Grissom et al. (2017) find that when at least 20% of the teachers and staff are race matched to their students, there is a significantly larger share of Black and Latino students in gifted programs.

Finally, students may seek informal mentorship from faculty in their school outside of their classrooms. Such mentorships may be formed through participation in extracurricular activities, such as after-school clubs or sports, or through simply being members of the same community. As previously mentioned, students of color may feel more comfortable working with teachers of color, as they are more likely to share similar backgrounds. Faculty of color may act as cultural translators, or an advisor on how to successfully navigate through schools or other institutions (Irvine, 1989; King, 1993).

Because we cannot observe student-faculty interactions outside of the teacher-student link within a classroom, we are not able disentangle the potential mechanisms through which race matched faculty impact a student's test scores with our data. While we detail each phenomenon separately, we refer to this joint phenomenon as the *faculty race match effect* for the remainder of the paper.

## 2.4 Data

This study uses administrative panel data from the Los Angeles Unified School District (LAUSD) spanning school years 2007-2008 through 2017-2018, which covers the start of the increase in the share of Latino teachers within the district. For each student, the data contain information on race,<sup>1</sup> school and grade, gender, English language learner (ELL) status, free- or reduced-priced lunch (FRL) eligibility, a student with disability (SWD) indicator, and standardized math and English language arts (ELA) test scores. These data link students to teachers and contain teacher characteristics such as race, grade taught, gender, years of experience, whether they have a Master's degree, and certification status (e.g., preliminary or fully certified). Each student-teacher pairing includes the course name and department (e.g., Algebra and Math), which we use to generate an indicator variable for math and ELA courses. Since test scores are only reported for grades 3

---

<sup>1</sup>Student and teacher race categories include: Asian, Black, Latino, and White. Filipino, Native American, Pacific Islander, and Multi-Race students have been subsumed into an "Other" category due to lack of sample size.

through 8, our analytical sample is constrained to these grades.

Table 2.1 summarizes the student and teacher characteristics by race. In the analytical sample, over 42% of the teaching force is Latino—one of the highest proportions throughout the country. While more work remains to achieve parity (over 77% of the LAUSD student population are Latino), these statistics indicate that Latino students have a higher chance of being race matched in LAUSD than in most other school districts. Although the teacher characteristics are fairly homogeneous across races, Latino teachers have fewer years of experience and are less likely to hold a Master’s degree or PhD. This can be explained by the recent trend of LAUSD hiring more Latino teachers.

Even though Latino teachers and students make up the plurality within their respective groups, White students match their teachers’ race at a much higher rate (62.7% compared to 39.3% for Latino students). On average, even though Latino teachers make up the largest proportion of the district by far, White students are still enrolled in schools and grades where the majority of their faculty are also White (60.9%). By comparing the remaining races from the teacher composition to the share of faculty for a given race, we find that same-race students and teachers tend to sort into the same schools and classrooms. For example, although Asian teachers make up 9.30% of the teacher racial distribution, on average, Asian students share the race of 13.3% of the faculty in their school and grade. We describe how we mitigate the concern for this form of sorting bias from our estimates in Section 4.2. Table 2.1 also shows that over a third (35.6%) of Latino students and over a quarter (27.3%) of Asian students in our sample are identified as English language learners. Also, almost ninety percent (89.3%) of Latino students and over three-quarters (76.4%) of Black students are eligible for free- or reduced-price lunches, which is a proxy for low-income students. Since gaps in test scores exist between low income/wealthier and ELL/English-fluent students in California (EdSource, 2017), it is especially important to account for both of these measures to avoid an omitted variable bias in the race match estimates. Our model is defined and described below.

**Table 2.1** Student & Teacher Summary Statistics by Race

Panel A Teacher-Year	(1) Asian	(2) Black	(3) Latino	(4) White	(5) Other	(6) Total
Female	0.867 (0.339)	0.837 (0.369)	0.788 (0.408)	0.785 (0.41)	0.781 (0.413)	0.799 (0.4)
Years of Experience	8.760 (2.414)	9.152 (2.098)	8.875 (2.307)	9.089 (2.116)	8.899 (2.293)	8.968 (2.235)
Master's or PhD	0.400 (0.49)	0.490 (0.5)	0.304 (0.46)	0.363 (0.481)	0.351 (0.477)	0.354 (0.478)
Fully Certified	0.917 (0.276)	0.907 (0.291)	0.913 (0.281)	0.929 (0.256)	0.907 (0.291)	0.918 (0.274)
<i>N</i>	15,113	16,374	68,672	56,932	5,450	162,541
% of Sample	9.30%	10.07%	42.25%	35.03%	3.35%	100.00%
Panel B Student-Year	(1) Asian	(2) Black	(3) Latino	(4) White	(5) Other	(6) Total
Race Matched	0.168 (0.374)	0.312 (0.463)	0.393 (0.488)	0.629 (0.483)	0.060 (0.238)	0.389 (0.488)
Faculty Race Match	0.133 (0.208)	0.269 (0.302)	0.416 (0.293)	0.609 (0.292)	0.058 (0.128)	0.401 (0.307)
Female	0.498 (0.5)	0.517 (0.5)	0.501 (0.5)	0.486 (0.5)	0.488 (0.5)	0.501 (0.5)
SWD	0.033 (0.178)	0.101 (0.302)	0.077 (0.266)	0.083 (0.276)	0.043 (0.202)	0.076 (0.265)
FRPL	0.530 (0.499)	0.764 (0.425)	0.893 (0.309)	0.334 (0.472)	0.586 (0.493)	0.811 (0.392)
ELL	0.273 (0.445)	0.015 (0.123)	0.356 (0.479)	0.147 (0.354)	0.193 (0.395)	0.307 (0.461)
<i>N</i>	56,167	89,064	1,051,438	120,805	41,814	1,359,288
% of Sample	4.13%	6.55%	77.35%	8.89%	3.08%	100.00%
Observations at the teacher-year and student-year level for school years 2007-08 through 2017-18. Standard deviations shown in parentheses. <i>SWD</i> = student with disability; <i>FRPL</i> = free- or reduced-price lunch; <i>ELL</i> = English language learner.						

## 2.5 Empirical Strategy

### 2.5.1 Model

We calculate the overall teacher-student and faculty-student race match effects under the following specification:

$$Y_{ijgst}^{sub} = \beta_0 + \beta_{RM}RM_{igst} + \beta_{FRM}\%FRM_{sgt} + \beta_{int}(RM_{igst} * \%FRM_{sgt}) \\ + \mathbf{X}_{igst}\Gamma_1 + \mathbf{Z}_{jgst}\Gamma_2 + \bar{\mathbf{Z}}_{j'gst}\Gamma_3 + \theta_t + \theta_i + \theta_{gs} + \epsilon_{ijgst}, \quad (2.1)$$

where  $i, j, g, s$ , and  $t$  indicate the student, teacher, grade, school, and year, respectively.  $Y$  represents the student's test score for subject  $sub \in \{math, ELA\}$ .  $RM$  is an indicator for whether the student is in a classroom with a race matched teacher.  $\%FRM$  represents the share of faculty that are the same race as the student within the grade-school, excluding their teacher.  $\beta_{int}$  is the effect of the interaction between teacher-student and faculty-student race matches and can be interpreted as the marginal impact of direct teacher-student race matching given an increase in the share of same race faculty. Year fixed effects ( $\theta_t$ ) control for any shock that could occur in a given year (e.g., the state exam is particularly difficult for one year).  $\mathbf{X}_{igst}$  and  $\mathbf{Z}_{jgst}$  represent vectors of student and teacher characteristics, respectively, presented in Table 2.1. Similarly,  $\bar{\mathbf{Z}}_{j'gst}$  represents an average of the teacher characteristics for the faculty members that share a grade-school with student  $i$ .<sup>2</sup>  $\theta_i$  and  $\theta_{gs}$  are the student and school-grade fixed effects and serve as the anchors to identifying the direct and faculty race match effects. The importance and identifying assumptions for these fixed effects are described in the next two subsections.

We also analyze how the direct teacher and faculty race match effects differ by race. To do this, specify a similar model that includes an indicator for the student's race,  $stuRace_i$ :

$$Y_{ijgst}^{sub} = \beta_0 + [\beta'_{RM}RM_{igst} + \beta'_{FRM}\%FRM_{sgt} + \beta'_{int}(RM_{igst} * \%FRM_{sgt})] * stuRace_i \\ + \mathbf{X}_{igst}\Gamma_1 + \mathbf{Z}_{jgst}\Gamma_2 + \bar{\mathbf{Z}}_{j'gst}\Gamma_3 + \theta_t + \theta_i + \theta_{gs} + \epsilon_{ijgst}, \quad (2.2)$$

where each component is defined in the same manner as in Equation 2.1. The race match coefficients

---

<sup>2</sup>The  $j'$  notation indicates all teachers except teacher  $j$ .



are interpreted similarly as before but now for each racial group.

### **2.5.2 Endogeneity of Faculty-Student Race Matching**

Teachers of the same race tend to teach in the same schools (e.g., Jackson and Bruegmann, 2009; Hanushek, 2004; Boyd et al., 2003) which, if left unaccounted for, could bias teacher-student and faculty-student race match effects. For instance, students may sort into schools with higher proportions of race matched teachers, thereby increasing the likelihood of being placed in a race matched classroom. For example, a Black student (or their parents) may choose to enroll in a school with higher proportions of black teachers to avoid the anxiety that comes with stereotype threat. In turn, the lower level of anxiety that student faces could improve test scores. While reducing stereotype threat is a potential mechanism for race match effects, the issue in not accounting for this type of sorting is that there could be unobserved characteristics (correlated with test scores) for the types of students that choose to avoid stereotype threat by sorting into schools with higher shares of same-raced faculty. To address this potential bias, we utilize an identification strategy similar to Chetty et al. (2014b), who use teacher turnovers to examine the effect of teacher value-added on long-run student outcomes. Specifically, by including a school-grade fixed effect, we focus on the changes to faculty within a grade-school cell, the faculty race match effect is identified off of the variation in the racial distribution of faculty within schools and grades across years. Examining the changes at the grade-school level ignores any racial sorting that occurs across schools. Though teachers are more likely to interact within a grade, teachers may also interact across grades. For this reason, the estimated faculty race match effect should be interpreted as a lower bound. Under the reasonable assumptions that 1) students do not choose to “fail” or “skip” grades because of the racial composition of teachers in the next grade and 2) teachers do not switch grades due to the racial composition of students, the faculty-student race match effect is identified.

### **2.5.3 Endogeneity of Student-Teacher Race Matching**

Although the school-grade fixed effect accounts for sorting across schools, it does not alleviate the concern that students and teachers may sort by race within schools. For example, students and teachers may also track into different courses disproportionately by race. If a certain race is more likely to track into higher ability courses, presumably with higher averaged test scores, there would exist an upward bias for the student-teacher race match effect. In order to account for this, we follow the example of previous literature (e.g., Egalite et al. (2015)) and include a student fixed effect within our main specification. This controls for any time-invariant, unobserved student characteristics, such as actively wanting to avoid stereotype threat or consistently being enrolled into tracked classrooms, from our teacher-student race match estimate. Additionally, in section 5.3, we use other proxies for student tracking by also controlling for a student's previous test score and the average test score for a student's classmates from the previous year.

## **2.6 Results**

### **2.6.1 Overall direct and faculty race match effects**

Table 2.2 shows the progression of results where we start with only a direct student-teacher race match and build up to Equation 1.1 for both math and ELA classes. In columns (1) and (4), we find large and significant impacts for the direct student-teacher race matches across both subjects, though we find a more prominent effect for math courses. We estimate that a student directly race matched with their teacher is expected to improve their math (ELA) test score by 0.019 (0.007) standard deviations, both of which are statistically significant at the 99.9% confidence level. While our results are somewhat larger than the findings in Egalite et al. (2015), the authors also find a more significant direct student-teacher race match effect for math (0.008 sd) than for ELA (0.002 sd) courses. Columns (2) and (5) show the estimates for only the share of race matched faculty. For example, increasing the share of race matched faculty (i.e., teachers within the same school and grade but not in the same classroom as a student) by one is expected to increase a student's

standardized ELA test score by 0.01 sd. Though a more reasonable interpretation, since the presented effects are linear, is that a 10 percentage point increase in faculty-student race matching impacts ELA test scores by 0.001 standard deviations. Finally, Columns (3) and (6) show the estimates for our main specification. Again, the coefficients for math courses are more prominent than for ELA with the direct race match estimate being 0.013 sd. We find that for a student not directly race matched to their teacher, a 10 percentage point increase in the share of race matched faculty is expected to decrease math test scores by 0.0009 sd. However, if we assume the same 10 percentage point increase in race matched faculty, a directly race matched student is expected to increase their math test score by  $0.013 - 0.0009 + 0.0014 = 0.0135$  standard deviations. Given the results found in previous direct race match studies, an additional improvement by 0.0005 sd is not negligible. For ELA, we estimate a positive faculty race match (assuming a 10 percentage point increase) effect of 0.0005 sd and 0.0046 sd for non-directly and directly race matched students, respectively. While we cannot empirically explain the differences found between the faculty race match effects across the subjects, one possible explanation is that math as a subject may be more objective, with a general consensus on content, sequence, and pedagogy than for ELA (Sun et al., 2017). This room for subjectivity may be more conducive to a teacher's ability to influence another teacher.

**Table 2.2** Overall Direct and Faculty Race Match Effects on Test Scores

	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
RM	0.019*** (0.002)		0.013*** (0.002)	0.007*** (0.002)		0.003 (0.002)
Share FRM		-0.005 (0.003)	-0.009** (0.003)		0.010*** (0.002)	0.005* (0.003)
RM*Share FRM			0.014*** (0.004)			0.011** (0.004)
N	1,277,361	1,277,361	1,277,361	1,359,288	1,359,288	1,359,288

Student-clustered standard errors shown in parentheses. All columns include school-grade, student, and year fixed effects. *RM* is an indicator variable for a direct student-teacher race match. *Share FRM* is the share of race match faculty at the school-grade level. *ELA* = English language arts.

When taken independently, the faculty race match effects may seem insignificant. However, comparing Columns (1) and (3) for math and Columns (4) and (6) for ELA indicates that excluding the share of faculty that are racially congruent to the student biases the direct race match effect. That is, ignoring the possibility for teacher collaboration, potential role model effects, or school culture inflates the student-teacher race match effect on test scores. The 0.0055 sd difference (0.019-0.0135) for math courses is fairly large given the magnitude found in previous race match studies. A similar exercise for the ELA results indicates that, estimating a direct teacher-student race match effect without including faculty race matches significantly biases the results.

### **2.6.2 Direct and faculty race match effects by student race**

In Table 2.3 we examine how the direct and faculty race match effects differ across student race. With an exception for Latino students, we do not find consistent direct and faculty race match effects. For Asian students, our results indicate there is no direct race match effect for either subject. However, we find a persistent negative effect for faculty race match in math and ELA courses. Similar for the Latino population in Florida, the Asian population within Los Angeles is culturally diverse with 30% of Asians identifying as having Chinese origin, 24% with Filipino origin, and 13% with Korean origin (Los Angeles Almanac, 2021). While on paper, Chinese and Filipino students and teachers are identified as being race matched, these cultures are very different. These cultural differences could be a driving factor in the negative results found. In line with the existing literature, we find that Black students are expected to benefit from a direct race match in math courses. Our results (0.025 sd) are quite similar to those found in Egalite et al. (2015) (0.018 sd). Our most consistent effects are for Latino students. Similar to the overall race match effects, we find that Latino students experience positive direct race match effects for both subjects, while experiencing negative (positive) faculty race match effects for math (ELA) courses. In the context of LAUSD, this result may be the most promising since the majority of students are Latino. A directly race matched Latino student with a 10 percentage point increase in race matched faculty is expected to improve on their test scores by 0.027 sd for math and 0.018 sd for ELA courses.

**Table 2.3** Direct and Faculty Race Match Effects on Test Scores by Student Race

	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Asian*RM	-0.002 (0.008)		0.004 (0.010)	-0.006 (0.007)		-0.008 (0.008)
Asian*Share FRM		-0.031* (0.014)	-0.022 (0.016)		-0.063*** (0.013)	-0.066*** (0.014)
Asian*RM*Share FRM			-0.034 (0.028)			0.014 (0.028)
Black*RM	0.015** (0.006)		0.025** (0.008)	-0.006 (0.005)		-0.004 (0.007)
Black*Share FRM		-0.013 (0.009)	-0.002 (0.011)		-0.002 (0.009)	0.001 (0.011)
Black*RM*Share FRM			-0.026 (0.015)			-0.007 (0.015)
Latino*RM	0.034*** (0.003)		0.026*** (0.004)	0.019*** (0.003)		0.016*** (0.004)
Latino*Share FRM		-0.004 (0.003)	-0.012** (0.004)		0.014*** (0.003)	0.011*** (0.003)
Latino*RM*Share FRM			0.020*** (0.004)			0.008 (0.004)
White*RM	0.003 (0.004)		-0.012 (0.008)	0.000 (0.004)		-0.024** (0.008)
White*Share FRM		0.007 (0.007)	-0.008 (0.010)		0.015* (0.006)	-0.009 (0.009)
White*RM*Share FRM			0.028* (0.013)			0.041*** (0.012)
N	1,277,361	1,277,361	1,277,361	1,359,288	1,359,288	1,359,288

Student-clustered standard errors shown in parentheses. All columns include school-grade, student, and year fixed effects. The Other race category has been suppressed from this table. *RM* is an indicator variable for a direct student-teacher race match. *Share FRM* is the share of race match faculty at the school-grade level. *ELA* = English language arts.

## **2.7 Robustness Checks**

### **2.7.1 Additional Controls for Student Tracking**

The timing for student tracking may not occur the same way across all of the grades in our sample or across subjects. For example, in ELA courses, students may be tracked based on perceived ability as early as Kindergarten (Brookings Institute, 2013). However, tracking for math courses is not typically done until middle school (Dreeben & Barr, 1988). As previously mentioned, students and teachers of the same race may track into courses, which could bias the race match estimates. To mitigate this concern, we include two additional proxies for tracking to the student fixed effects into the model specification. The first measure is a student's test score from the previous year. If students are able to vastly improve in a particular subject from one grade to the next, then the student fixed effect, which only controls for time-invariant unobservables, would no longer be a sufficient way to control for student tracking. Presumably, a student who scores very well on the standardized test for a given subject will be placed into a tracked classroom in the following year for that same subject. This control allows for a student's ability to vary across time, while still taking into account the possibility for student tracking. For the second measure, we use the average test scores for a student's classmates in the previous year. If this average is particularly high, that student may be more likely to track into an advanced class for the given subject in the following year. Table 2.4 shows that our results are nearly identical when including either of the proxies for student tracking, which indicates that the student fixed effect approach may sufficiently account for student tracking.

**Table 2.4** Race match effects with additional controls for student tracking

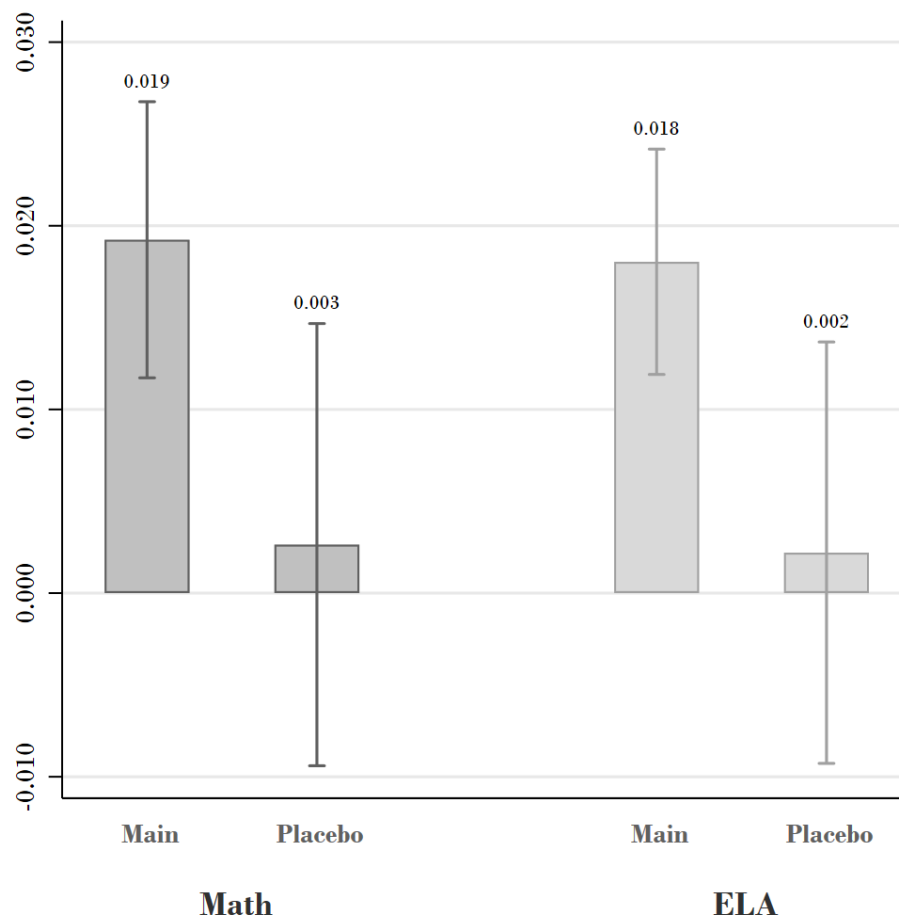
	Math			ELA		
	(1) Main	(2) Pre-Test	(3) Class Pre-Test	(4) Main	(5) Pre-Test	(6) Class Pre-Test
Asian*RM	0.004 (0.010)	0.007 (0.010)	0.003 (0.010)	-0.008 (0.008)	-0.010 (0.008)	-0.009 (0.008)
Asian*Share FRM	-0.022 (0.016)	-0.014 (0.016)	-0.026 (0.016)	-0.066*** (0.014)	-0.067*** (0.014)	-0.065*** (0.014)
Asian*RM*Share FRM	-0.034 (0.028)	-0.048 (0.028)	-0.037 (0.028)	0.014 (0.028)	0.018 (0.028)	0.008 (0.028)
Black*RM	0.025** (0.008)	0.024** (0.008)	0.027*** (0.008)	-0.004 (0.007)	-0.005 (0.007)	-0.004 (0.007)
Black*Share FRM	-0.002 (0.011)	-0.001 (0.011)	-0.001 (0.011)	0.001 (0.011)	0.001 (0.011)	0.001 (0.011)
Black*RM*Share FRM	-0.026 (0.015)	-0.024 (0.015)	-0.027 (0.015)	-0.007 (0.015)	-0.008 (0.015)	-0.005 (0.015)
Latino*RM	0.026*** (0.004)	0.025*** (0.004)	0.026*** (0.004)	0.016*** (0.004)	0.016*** (0.004)	0.018*** (0.004)
Latino*Share FRM	-0.012** (0.004)	-0.012** (0.004)	-0.015*** (0.004)	0.011*** (0.003)	0.012*** (0.003)	0.008* (0.003)
Latino*RM*Share FRM	0.020*** (0.004)	0.019*** (0.004)	0.020*** (0.004)	0.008 (0.004)	0.007 (0.004)	0.006 (0.004)
White*RM	-0.012 (0.008)	-0.010 (0.008)	-0.014 (0.008)	-0.024** (0.008)	-0.022** (0.008)	-0.024** (0.008)
White*Share FRM	-0.008 (0.010)	-0.007 (0.010)	-0.012 (0.010)	-0.009 (0.009)	-0.008 (0.009)	-0.009 (0.009)
White*RM*Share FRM	0.028* (0.013)	0.023 (0.012)	0.031* (0.013)	0.041*** (0.012)	0.038*** (0.011)	0.043*** (0.012)
N	1,277,361	1,277,361	1,277,361	1,359,288	1,359,288	1,359,288
Student-clustered standard errors shown in parentheses. All columns include school-grade, student, and year fixed effects. The Other race category has been suppressed from this table. <i>RM</i> is an indicator variable for a direct student-teacher race match. <i>Share FRM</i> is the share of race match faculty at the school-grade level. <i>ELA</i> = English language arts.						

## 2.7.2 Placebo Test

In theory, a student's outcome should not be impacted by their race match status in the following year. Figure 2.1 shows the results of a placebo test using the share of faculty race match and the

direct race match for the following year to simulate a placebo. To ensure a proper comparison, the same sample is used for the main regression model and the placebo test. The figure illustrates the sum of the race match components (i.e., direct race match, share faculty race match, and the interaction) with the adjusted 95% confidence interval. For both subjects, the coefficient for the placebo test is not statistically different from zero, which provides some evidence for the validity of the main results.

**Figure 2.1** Placebo Test using following Year Race Match Components



The first bar represent the sum of coefficients for  $ShareFRM_{igst}$ ,  $RM_{igst}$ , and the interaction term from equation 2.1. The second bar estimates a placebo test by replacing each race match component with the following year's race match component. The same sample was used in both regressions with  $N=1,277,361$  for math and  $N=1,359,288$  for ELA. The coefficients are interpreted as a change in standard deviations. The whiskers represent a 95% confidence interval around the estimated effect.



## 2.8 Race Match by Elementary and Middle School Grades

In Table 2.5, we further explore the effects for Latino students by separating the sample into elementary (grades 3 through 5) and middle (grades 6 through 8) schools. For both subjects, the direct race match effect is larger in earlier grades than for middle school grades. However, the effect is still substantial across both subsamples. The impact of same race faculty seems to be driven by students in middle schools.

**Table 2.5** Race Match Effects on Test Scores by Elementary/Middle School for Latino Students

	Math			ELA		
	(1) All	(2) Elem	(3) Mid	(4) All	(5) Elem	(6) Mid
Latino*RM	0.026*** (0.004)	0.032*** (0.006)	0.020** (0.006)	0.016*** (0.004)	0.019** (0.006)	0.014* (0.006)
Latino*%FRM	-0.012** (0.004)	0.003 (0.006)	-0.018** (0.007)	0.011*** (0.003)	0.005 (0.005)	0.018** (0.006)
Latino*RM*%FRM	0.020*** (0.004)	0.000 (0.007)	0.035*** (0.008)	0.008 (0.004)	0.004 (0.006)	0.006 (0.008)
N	1,277,361	569,530	561,657	1,359,288	571,020	651,121

Student-clustered standard errors shown in parentheses. All columns include school-grade, student, and year fixed effects. Columns (1) and (4) show the coefficients from the main specification. Columns (2) and (5) restrict the sample to students in grades 3 through 5. Columns (3) and (6) restrict the sample to students in grades 6 through 8. The all other race categories have been suppressed from the table. *RM* is an indicator variable for a direct student-teacher race match. *Share FRM* is the share of race match faculty at the school-grade level. *ELA* = English language arts.

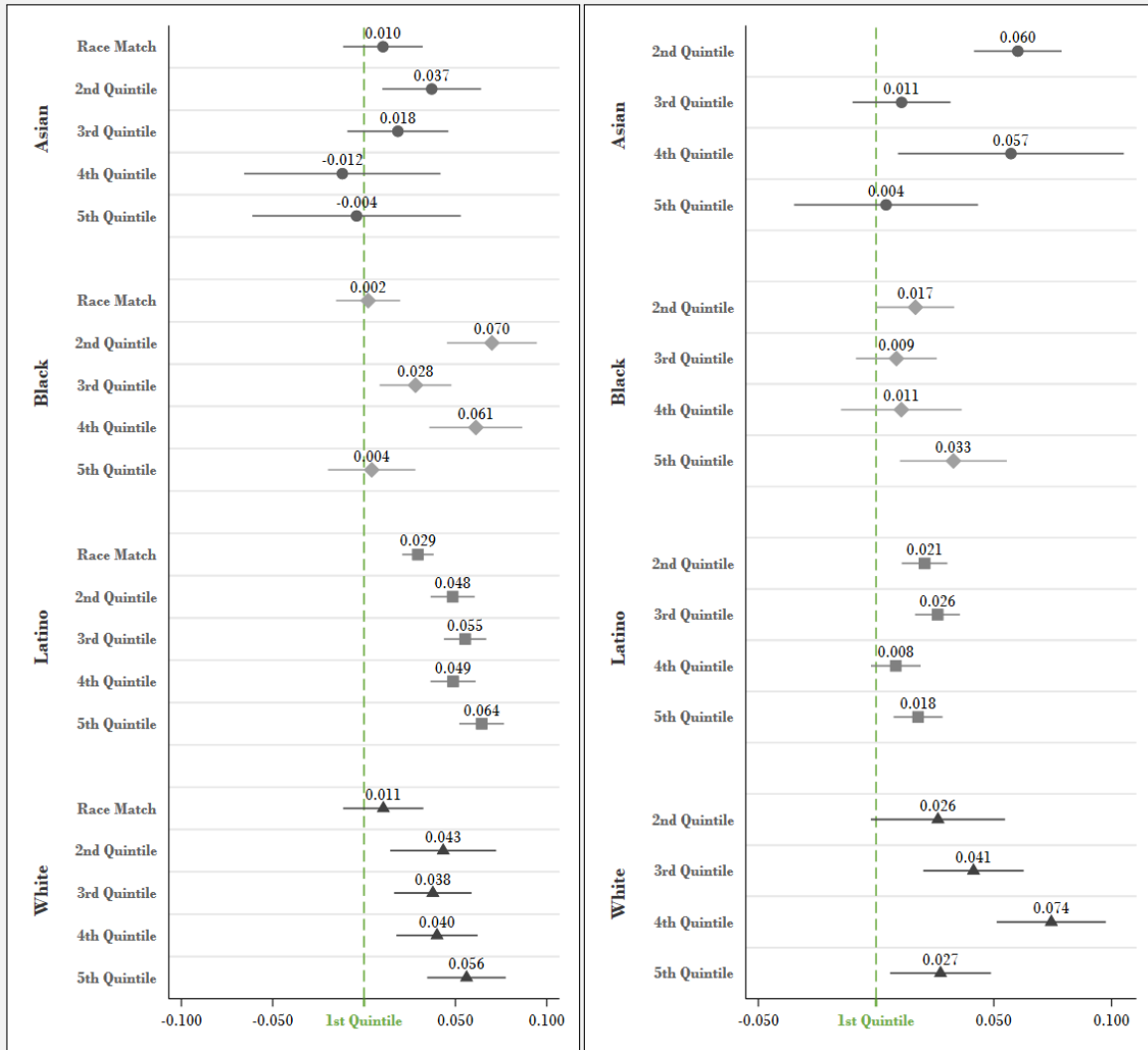
### 2.8.1 Faculty Race Match Quartiles

The race match effect may differ based on a student's current exposure to similarly raced teachers. The impact of race matching may be larger for students with lower shares of race matched teachers and diminish as this share increases. Figures 2.2 (math) and 2.3 (ELA) illustrate the non-linear race match effects by grouping the share of faculty race match into quintiles. For both figures, Panel A shows the sum of each race match component (i.e., direct, faculty, and the interaction) for race matched students. For each race, the top point represents directly race matched students within the first quintile (i.e., lowest share of race matched faculty). Each subsequent point represents race matched students within the given quintile of race matched faculty. In general, there do not seem to be any non-linear effects in the share FRM for directly race matched students. One exception is for directly race matched Latino students, where there seems to be a constant effect of  $0.035sd$  beyond the first quintile of share FRM. Panel B shows analogous results for share FRM for students without a direct race match. In general, there do not seem to be any non-linear effects for these students, though is partially due to the noise of the estimates.

**Figure 2.2 Math Test Scores: Effects by Quintile of Share Faculty Race Match**

**Panel A: Direct Race Match**

**Panel B: No Direct Race Match**



Test scores are standardized at the grade-year level. Students are binned into each quintile based on the distribution of share faculty race match (FRM). The first quintile represents students with no to low shares of race matched faculty, and the fifth quintile represents students with higher shares of race matched faculty. Panel A shows the results for students with a direct race match, where each point represents the linear combination for the coefficients of the respective FRM quintile, the direct race match, and their interaction. Panel B shows the results for students without a direct race match, each point represents the quintile of share FRM. The whiskers represent a 95% confidence interval. The baseline (represented by the dotted line) group are students without a direct race match and within the first quintile of share faculty race match.

**Figure 2.3 ELA Test Scores: Effects by Quintile of Share Faculty Race Match**



Test scores are standardized at the grade-year level. Students are binned into each quintile based on the distribution of share faculty race match (FRM). The first quintile represents students with no to low shares of race matched faculty, and the fifth quintile represents students with higher shares of race matched faculty. Panel A shows the results for students with a direct race match, where each point represents the linear combination for the coefficients of the respective FRM quintile, the direct race match, and their interaction. Panel B shows the results for students without a direct race match, each point represents the quintile of share FRM. The whiskers represent a 95% confidence interval. The baseline (represented by the dotted line) group are students without a direct race match and within the first quintile of share faculty race match.

## 2.9 Conclusion

Race match studies suggest that shrinking the disparity between the racial distributions of teachers and students might be an important strategy to help close the racial achievement gap. However,

this does not necessarily mean that all students of color need to have a teacher of color. Overall, we find that faculty members of congruent races can also have a significant and positive impacts on students. While exploring the exact mechanisms is beyond the scope of this paper, a potential driver for these effects could be that teachers are able to learn how to improve their instruction for students of color when a teacher of color is present.

Moreover, because race match effects differ across races (e.g., Egalite et al., 2015), it is not enough to only understand these effects for Black and White students. By studying how race matched teachers and faculty impact test scores in Los Angeles, an urban setting with a large Latino population, we are able to fill this gap within the literature. Our findings indicate that matching Latino students to Latino teachers improve test scores for math and ELA, while other Latino faculty improve test scores in these students' ELA courses. The direct student-teacher results are similar for Latino students in elementary grades as for those in middle school grades. However, the faculty race match effect is only present in middle school grades. Also, both race match effects are robust to various measures of student tracking. Increasing the labor supply of Latino teachers may provide as a crucial catalyst in decreasing the achievement gaps found between Latino and White students. However, we do not find consistent direct or faculty race match effects for other students of color. Two exceptions are that race matched Black students seem to benefit for math courses and that Asian students see lower test scores with increases in race matched faculty. However, the diversity within the Asian population in Los Angeles could be a major driving factor in these results.

While the faculty race match effects are generally null, they are still important to include into models for direct race match effects. Omitting these measures bias the student-teacher race match effect. If researchers and policymakers truly intend to diversify the teacher pool as a potential way to shrink achievement gaps, then it is important to correctly measure the impact that teachers of color have for students of color.

## CHAPTER 3

### ARE EFFECTIVE TEACHERS FOR STUDENTS WITH DISABILITIES EFFECTIVE TEACHERS FOR ALL?

#### 3.1 Disclaimer

This chapter was co-authored with Ijun Lai (ILai@mathematica-mpr.com), Neil Filosa (filosane@msu.edu), Scott Imberman (imberman@msu.edu), Nathan Jones (ndjones@bu.edu), and Katharine Strunk (kstrunk@msu.edu). All authors have approved that this work be included as a chapter in my dissertation.

#### 3.2 Introduction

Research measuring teacher quality has made tremendous strides over the last several years, but much of this work has shared a common assumption—that teachers’ effectiveness does not vary across student subgroups. Student growth measures and common evaluation metrics like in-class observations tend to focus on the overall class; an effective teacher is one who generates higher levels of academic growth, on average, and one who uses practices that reflect a single definition of good teaching.<sup>1</sup> But such approaches obscure the reality that teachers may have skills that benefit some groups of students more than others. Indeed, Loeb et al. (2014) and Master et al. (2016) have documented that teachers are differentially effective at improving academic outcomes for English Learners (ELs), and that prior experience and training are both predictive of teachers’ success in supporting these students’ academic growth. The work on ELs documents that without disaggregating teacher effectiveness below the class-level, we lack key information about whether teachers are sufficiently meeting the needs of critical student subpopulations. Two broad questions emerge from this concern. First, are some teachers effective at supporting some students more than others? Second, if so, do schools assign students to teachers who are best able to meet their needs?

---

<sup>1</sup>Research has documented that observation tools commonly used in general education may not capture practices known to support students with disabilities (Jones et al., 2022; Morris-Mathews et al., 2021).

This paper focuses on whether teachers are differentially effective and their assignment to students with disabilities (SWDs). This student population is important for a few key reasons. For one, SWDs make up approximately 13% of the K-12 student population. The majority of SWDs receive most of their instruction in the general education class.<sup>2</sup> Thus, it is important to understand how general educators impact SWDs. Second, SWDs tend to score lower on math and reading achievement assessments than their nondisabled peers (e.g., Chudowsky and Chudowsky, 2009; Schulte et al., 2016) and many experience worse long-term academic outcomes (Stiefel et al., 2021). Understanding if there are teachers who are more effective with SWDs may help school and district leaders pair SWDs with teachers who can best serve them.

Why might we expect teachers to differ in their effectiveness with SWDs? There is ample literature suggesting that specific instructional practices – such as explicit, systematic instruction – are likely to improve academic outcomes for SWDs (e.g., Gersten et al., 2009a,b; Jones and Winters, 2022; Stockard et al., 2018; Torgesen et al., 2001). And these practices may benefit SWDs specifically relative to their nondisabled peers (Connor et al., 2011, 2018). Yet many general educators lack the necessary training to use practices known to support SWDs (e.g., Brownell et al., 2010; Cook, 2002; Sindelar et al., 2010). Only seven states require general education teachers to complete coursework for working with SWDs, and only two require clinical experiences working with these students (Galiatsos et al., 2019). Existing studies suggest that general education teachers receive minimal coverage of special education teaching methods in their coursework, and few have any practice opportunities focused on SWDs (Blanton et al., 2011, 2018; Florian, 2012; Galiatsos et al., 2019). Surveys of general education teachers routinely find that many of them feel underprepared to support the diverse educational needs of SWDs (e.g., Kamens et al., 2000; Sadler, 2005). Thus, it is likely the case that many SWDs are assigned to teachers who do not have the requisite expertise for supporting them. And, within a single classroom, teachers may be more or less effective meeting the needs of SWDs or non-SWDs.

To test the relative effectiveness of general education teachers with SWDs, we leverage Los

---

<sup>2</sup>Over 60% of SWDs spend 80% or more of their school time in general education (U.S. Department of Education, National Center for Education Statistics, 2018).

Angeles Unified School District (LAUSD) administrative data from SY 2007-2008 through SY 2017-2018. As the second largest school district in the United States, Los Angeles provides a useful context to study this question as it has diverse teacher and student populations and a large population of SWDs spread across teachers and schools. Using these data, we generate two different sets of value-added measures (VAMs) for each teacher: one focusing on teachers' effectiveness in improving test scores for SWDs (SWD VAMs) and one focusing on non-SWDs (non-SWD VAMs). With these two measures, we also explore whether teachers have a relative advantage in teaching SWDs, defined as being higher in the SWD VAM distribution than in the non-SWD VAM distribution.<sup>3</sup> We use all three measures to explore the following research questions in the Los Angeles context:

1. Is there evidence that a substantial portion of teachers have a relative advantage in teaching SWD vs non-SWDs and vice-versa?
2. Are higher SWD/non-SWD VAMs associated with observable teacher characteristics? If so, do teachers with higher SWD VAMs share the same observable characteristics as teachers with higher non-SWD VAMs?
3. Do SWDs sort into classes taught by a teacher with high SWD VAMs?
4. Are SWDs sorted to teachers with a relative advantage for instructing SWDs?
5. How do retention rates compare for teachers with high versus low SWD VAMs? Are schools more or less likely to retain teachers with relative advantages for instructing SWDs than those with relative disadvantages?

This study contributes to the literature in several ways. First, we provide multiple measures of teacher effectiveness and document whether teachers are differentially effective at instructing students with and without disabilities. The intuitive question of whether teachers are more effective

---

<sup>3</sup>For example, a teacher who is in the 30th percentile of the non-SWD VAM distribution but in the 45th percentile of the SWD VAM distribution would have a relative advantage in teaching SWDs even though she is below the median in both distributions (and hence is at an absolute disadvantage).



with one student population than with another brings with it several complex analytical challenges. We demonstrate novel approaches for how researchers could conduct similar kinds of analyses moving forward. Second, we expand on previous literature that has sought to link observable teacher characteristics with more effective teachers (e.g., Kane and Staiger, 2008); here, we conduct a similar analysis on observable teacher characteristics and teachers' SWD and non-SWD VAMs. Finally, we explore how student placement and teacher mobility are correlated with teachers' efficacy and relative advantage. Placing SWDs in classrooms with teachers who are the most effective at instructing SWDs can potentially raise the test scores of these students, but only if schools are able to retain these teachers.

### **3.3 Data**

#### **3.3.1 Background and Context**

We use data from Los Angeles Unified School District (LAUSD), the second-largest school system in the United States. In 2021, LAUSD enrolled over 520,000 students in grades K-12 and employed nearly 24,000 K-12 teachers (LAUSD, 2022). The LAUSD administrative data contain detailed information on student and teacher characteristics, and the sample is highly diverse across student, teacher, school, and community characteristics including socioeconomic status, race/ethnicity, disability status, teacher quality, and achievement.

In LAUSD, approximately 10% of students are identified as having a disability, and 60% of students with mild or moderate disabilities spend most of their day in general education classrooms (Swaak, 2020). School personnel, outside professionals, and an SWD's parents place the student in the most appropriate educational setting based on their individualized needs. However, to the maximum extent appropriate, LAUSD requires that students spend time with non-disabled peers in a general education classroom setting (LAUSD, 2018). Under this condition, it may be preferable for principals to pre-identify which teachers have an advantage in teaching SWDs to help make classroom assignments. In the following section, we describe the required data structure for generating SWD and non-SWD VAMs.

### **3.3.2 Sample and Data Requirements for Generating SWD and non-SWD VAMs**

This study uses student- and teacher-level administrative data spanning from the academic years 2007-2008 through 2017–2018, provided by LAUSD’s Office of Data and Accountability and the Division of Human Resources. The data are at the student-year level and include demographic information such as disability status (detailed below), race/ethnicity, gender, free or reduced-price lunch (FRL) eligibility, and English Learner (EL) status, as well as math and English language arts (ELA) state test scores. We normalize each subject’s test scores to have a mean of zero and standard deviation of one for each grade-year combination. We include both math and ELA scores because of evidence that SWDs may have different challenges in each subject (e.g., Child et al., 2019; Fuchs et al., 2016). In LAUSD, test score data are collected for students in grades 3 through 8, which limits our sample to these grades. Additionally, while studying teacher quality for SWDs in all classrooms would be ideal, test scores (and correspondingly, VAM calculations) are only available for math and ELA courses.

In this study, we are interested in analyzing the differences in exposure to teacher quality for SWDs and non-SWDs. To compare teachers’ non-SWD VAMs relative to their SWD VAMs, our analytic dataset is limited to students attending mainstream classrooms. To ensure teachers have enough observations for both student groups, we further limit the analytical sample of teachers to those with at least 10 non-SWDs and 10 SWDs, totaling at least 20 student observations, between SY 2007-08 and 2017-18. This restriction limits the teacher sample to general education teachers with substantial experience in instructing SWDs. It is important to note that our analysis eliminates general education teachers who do not teach enough SWDs over the period of this study. Given the test score data and teacher restrictions, we observe 578 unique schools, 7,207 unique teachers, and 661,789 unique students with 10.7% of students having ever been classified as SWDs.

Table 3.1 illustrates the demographic and standardized test information for all students and teachers in our analytical sample. In the first row of Panel A, each student is represented only once. Since many student characteristics change over time (e.g., disability status, free or reduced-priced lunch status, English Learner status), the interpretation of student characteristics for this

row is the percent of students who have ever been classified for the specific characteristic. For example, 10.7% of the students within our sample were ever classified as SWDs during the sample period. The test score columns are first averaged within each student across years, then averaged again across students. The LAUSD data include information on 13 disability subcategories. We conduct analyses for SWDs overall and then, for illustrative purposes, focus on the three disability subcategories with the largest percentages of students; we subsume the remaining disability types into an “other” category. The disability categories in our study include Autism, SLD (specific learning disability), LS Imp. (language & speech impairment), and Other Dis. (other disability). These categories are not mutually exclusive, which explains why the sum of these columns is greater than the SWD column. Almost 85% of students have ever been eligible for free or reduced-priced lunch (FRL), 26.6% have ever been classified as an English Learner (EL), 73.8% of students are Latino/a, and only 10.3% of students are White.

**Table 3.1** Student and Teacher Descriptive Characteristics

Panel A: Student	SWD	Autism	SLD	LS Imp.	Other Disability	FRL	EL	Female	Asian	Black	Latino/a	White	Other Race	Math Test	ELA Test
<i>Student</i> <i>N</i> = 661,789	0.107 (0.309)	0.008 (0.092)	0.071 (0.257)	0.012 (0.110)	0.023 (0.150)	0.848 (0.359)	0.266 (0.442)	0.494 (0.500)	0.041 (0.198)	0.088 (0.283)	0.738 (0.440)	0.103 (0.304)	0.030 (0.170)	0.057 (0.966)	-0.006 (0.944)
<i>Student-Year</i> Non-SWD <i>N</i> = 1,465,535	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.785 (0.411)	0.182 (0.386)	0.511 (0.500)	0.043 (0.202)	0.074 (0.262)	0.745 (0.436)	0.107 (0.309)	0.032 (0.176)	0.126 (0.992)	0.078 (0.951)
SWD <i>N</i> = 158,373	1.000 (0.000)	0.081 (0.273)	0.679 (0.467)	0.094 (0.291)	0.202 (0.402)	0.809 (0.393)	0.418 (0.493)	0.349 (0.477)	0.019 (0.135)	0.096 (0.294)	0.756 (0.429)	0.111 (0.315)	0.018 (0.133)	-0.577 (0.852)	-0.692 (0.865)
Autism <i>N</i> = 12,877	1.000 (0.000)	1.000 (0.000)	0.058 (0.233)	0.018 (0.131)	0.049 (0.217)	0.646 (0.478)	0.222 (0.416)	0.144 (0.351)	0.066 (0.248)	0.078 (0.268)	0.562 (0.496)	0.251 (0.433)	0.043 (0.204)	-0.067 (1.030)	-0.157 (1.000)
SLD <i>N</i> = 107,512	1.000 (0.000)	0.007 (0.083)	1.000 (0.000)	0.026 (0.159)	0.036 (0.185)	0.848 (0.359)	0.498 (0.500)	0.395 (0.489)	0.011 (0.104)	0.086 (0.280)	0.817 (0.387)	0.073 (0.261)	0.013 (0.112)	-0.740 (0.740)	-0.878 (0.747)
LS Imp. <i>N</i> = 14,836	1.000 (0.000)	0.015 (0.122)	0.187 (0.390)	1.000 (0.000)	0.060 (0.238)	0.807 (0.395)	0.330 (0.470)	0.260 (0.439)	0.032 (0.175)	0.086 (0.280)	0.745 (0.436)	0.116 (0.320)	0.022 (0.146)	-0.103 (0.967)	-0.234 (0.967)
Other Disability <i>N</i> = 32,068	1.000 (0.000)	0.020 (0.140)	0.119 (0.324)	0.028 (0.164)	1.000 (0.000)	0.760 (0.427)	0.267 (0.443)	0.296 (0.457)	0.019 (0.137)	0.138 (0.345)	0.639 (0.480)	0.179 (0.384)	0.024 (0.153)	-0.472 (0.883)	-0.511 (0.901)
Panel B: Teacher															
	SpEd Credential	Master's or PhD	Experience					Female	Asian	Black	Latino/a	White	Other Race		
			Novice (1-2 yrs)	Early (3-5 yrs)	Middle (6-10 yrs)	Late (10+ yrs)									
<i>Teacher-Year</i> <i>N</i> = 45,139	0.015 (0.122)	0.364 (0.481)	0.022 (0.145)	0.063 (0.243)	0.150 (0.357)	0.766 (0.424)		0.694 (0.461)	0.087 (0.282)	0.097 (0.296)	0.390 (0.488)	0.392 (0.488)	0.034 (0.180)		

A student's status for a characteristic may change across years. The first row represents the percent of students that have ever been indicated with the respective status (except test scores, which represent means). Remaining rows indicate student-year observations pooled across SY 2007-08 to SY 2017-18. Sample contains students paired with teachers with at least 10 SWD and non-SWD observations across the dataset. SWD (student with disability). SLD (specific learning disability). LS Imp. (language & speech impairment). FRL (free or reduced-price lunch). EL (English learner). SpEd (special education).

The remaining rows for student characteristics in Table 1 are listed at the student-year level (i.e., the level of our analysis), and show how student characteristics vary by the different disability types. The second and third rows separate the overall student body between non-SWDs and SWDs. This separation highlights some interesting differences between the two groups of students and provides additional avenues where teachers may differ in effectiveness. Aside from the substantial difference in standardized test scores, SWDs are far more likely to be ELs and male. Of student-years classified as a SWD, 8.1% are classified under Autism, 67.9% under SLD, 9.4% under LS Imp., and 20.2% under Other Dis.

Table 3.1 Panel B describes the characteristics of teachers within our sample at the teacher-year level. 1.5% of observations hold a special education teaching credential, and 36.4% have a master's or PhD degree. Most teachers within our sample are relatively experienced, with only 2.2% representing teachers within their first two years of teaching (Novice), 6.3% with three to five years of experience (Early Career), 15.0% within six to nine years (Middle Career), and 76.6% with ten or more years of experience (Late Career). Note that since we are restricting the sample to teachers who have had at least ten SWD and ten non-SWDs in their career, this threshold skews our sample towards more experienced teachers than the LAUSD average.

We compare teacher characteristics from our analytic sample to the excluded sample in Table 3.2. Also, in Table 3.2, we explore changes to our 10 SWD and 10 non-SWD thresholds by examining how the teacher sample differs when we change the threshold to 6 or 15 students for each subgroup. Compared to the excluded sample, the share of teachers that hold a special education credential decreases substantially once we set the threshold at 6 or above. This is logical since teachers with special education credentials are likely to teach few non-SWDs, whom we need for estimating the non-SWD VAMs. Also, the share of novice and early-career teachers decrease when comparing the excluded sample to our analytic sample. This is likely due to teachers with more experience having more students, both SWDs and non-SWDs, across the 11 years of the dataset. As a robustness check, we also provide the results at each threshold level for *Research Question 1* in Figure G.1 and for *Research Question 3* in Table G.1 in Appendix G. Due to the similarity in

**Table 3.2** Share of Teachers by Descriptive Characteristics, Varying Minimum Threshold of Students

Threshold Sample	≤5 (Out)	≥6	≥10	≥15
SpEd Credential	19%	1%	1%	2%
Master's or PhD	38%	37%	36%	36%
Exp: Novice (1-2 yrs)	7%	2%	2%	2%
Exp: Early (3-5 yrs)	9%	6%	6%	6%
Exp: Mid (6-9 yrs)	16%	15%	15%	15%
Exp: Late (10+ yrs)	68%	76%	77%	77%
Female	73%	70%	69%	68%
Asian	10%	9%	9%	9%
Black	13%	10%	10%	9%
Latino/a	35%	39%	39%	39%
White	38%	39%	39%	40%
Other Race	4%	3%	3%	3%

Threshold samples contain teachers with at least the given threshold quantity (i.e., 6, 10, or 15) of SWD and non-SWD observations across the panel dataset. Data are at teacher-year level. Out of sample teacher-years teach 33% of student-year observations. SpEd (special education). Exp (years of experience). SWD (student with disability).

results across thresholds, we focus on the 10-threshold sample for the remainder of this paper.

### 3.3.3 SWD and non-SWD VAM Calculations

For each subject with available test score data (math and ELA) and year, we create two separate VAMs for each teacher to measure their contribution to student achievement growth, as measured by test scores, for SWDs and for non-SWDs. We estimate the following model separately by students' SWD status:

$$Ach_{ijst}^{subject} = \beta_1 Ach_{ijst-1}^{math} + \beta_2 Ach_{ijst-1}^{ela} + \mathbf{X}_{ijst} \mathbf{\Gamma} + \mathbf{T}_j \mathbf{\Omega} + \theta_t + \epsilon_{ijst}, \quad (3.1)$$

where  $Ach$  is either math or ELA achievement, standardized within subject and year, for student

$i$  with teacher  $j$  in year  $t$ . We use a leave-year-out method following Chetty et al. (2014a), which eliminates the concern of measurement error of previous years impacting the VAM estimate for year  $t$ .<sup>4</sup> Also, we control for students' achievement in the prior year in both math and ELA, as the inclusion of the second subject helps to mitigate bias due to sorting and to attenuate measurement error Lockwood and Mccaffrey (2014).  $\mathbf{X}$  is a vector of student demographic characteristics and is used to control for potential unobservable factors that may influence a student's test score outside of a teacher's control, which includes indicators of student race, gender, FRL eligibility, EL status, and grade level. Teachers' VAMs are estimated by the coefficients on a set of teacher fixed effects ( $\mathbf{T}$ ) and are normalized at the teacher-year level. Normalization allows us to interpret VAMs as the number of standard deviations from the mean teacher in terms of VAM score.  $\theta_t$  are year fixed effects, and  $\epsilon$  is a normally distributed error term. We estimate heteroskedastic-robust standard errors. This specification was chosen based on a detailed review of the current best practices in VAM modeling, summarized in Koedel et al. (2015).

### **3.4 Research Question 1. Is there evidence that a substantial portion of teachers have a relative advantage in teaching SWD vs non-SWDs and vice-versa?**

#### **3.4.1 RQ1 Methods**

We begin by exploring differences in teachers' relative ranks in SWD and non-SWD VAMs across LAUSD teachers and years. Our analysis focuses on rank ordering because the SWD and non-SWD VAMs are measured from different student populations making them ordinal but not cardinal comparable. That is, since VAMs are measured relative to the mean growth in each group, a SWD VAM of say, 2.3 does not reflect the same learning growth as a non-SWD VAM of the same value. Hence, we cannot say whether a teacher increases achievement in SWDs by a higher or lower amount than she increases achievement in non-SWDs. Instead, we can only say whether a teacher is higher or lower in the distribution of SWD performance amongst her peers relative to her position

---

<sup>4</sup>This model shrinks estimates using the autocovariance of mean test score residuals across years, which weighs more recent years more heavily compared to years that are further away from  $t$ .

in the distribution of non-SWD performance.

For each VAM (i.e., math SWD, math non-SWD, ELA SWD, and ELA non-SWD), we define the relative rank to be the percentile the teacher falls into within the VAM distribution for a given year. Once these values are assigned to each teacher, we compute the difference between the non-SWD and SWD VAM percentiles for each subject and year. We label this statistic as the difference in VAMs (DVAM), which is defined as:

$$DVAM_{jt}^{sub} = Ptile_{jt} \left( VAM_{Non-SWD}^{subject} \right) - Ptile_{jt} \left( VAM_{SWD}^{subject} \right), \quad (3.2)$$

where  $Ptile(.)$  converts its argument into a percentile. Using Equation 3.2, we define a teacher as having a relative advantage in instructing SWDs when their DVAM is negative. In other words, a teacher who falls in a higher percentile in the SWD VAM distribution than in the non-SWD VAM distribution is defined as having a relative advantage in instructing SWDs. As an example, a teacher in the 80th percentile for SWD VAM and the 50th percentile for non-SWD VAM would have a DVAM of -30 (i.e., 30 percentile points higher in the SWD VAM distribution than the non-SWD VAM distribution).

One concern regarding teacher DVAMs is that VAM estimates can be quite noisy, and so some of the within-teacher variation we see between the two types of VAMs is likely due to random error. To understand the potential extent of this error, we construct a new set of VAMs based on randomly generated groups of SWDs and non-SWDs. Specifically, we randomly assign SWD status to each student (while preserving the total size of and share of SWDs in a teacher's class) and calculate two separate VAMs by group (i.e., SWD or non-SWD) for each teacher using the same method as described above. Since disability status in this case is randomly assigned, any differences in a teacher's "random" SWD and non-SWD VAMs would be due to random noise. We repeat this process 100 times and, following Equation 3.2, we generate a randomized DVAM for math and another for ELA for each randomization. Finally, separating by subject, we pool each randomized DVAM to generate the random DVAM distribution. This exercise provides us with a baseline distribution to assess how much the original DVAM distribution differs from random



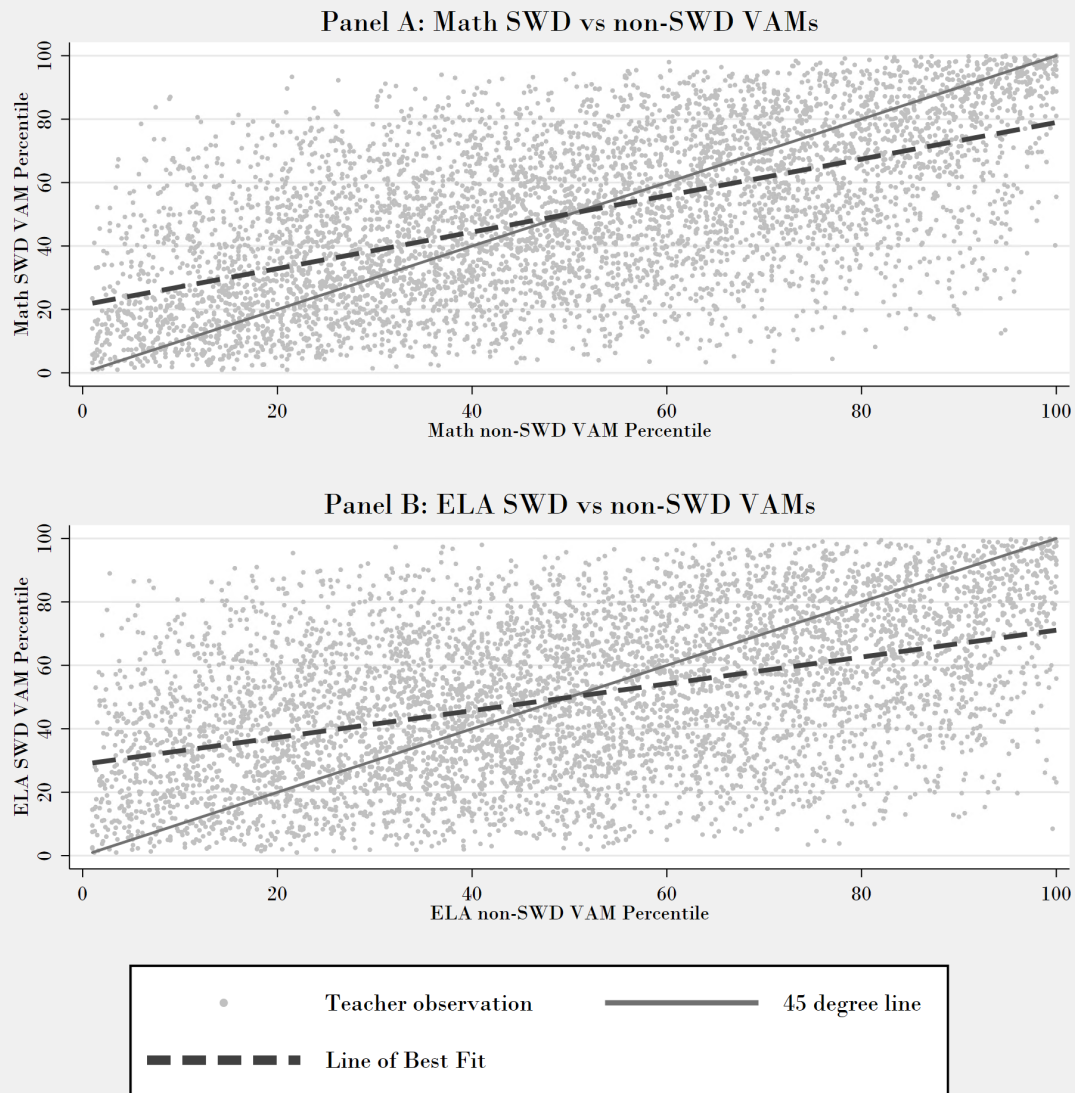
estimation error.

In theory, randomizing SWD status should generate SWD and non-SWD VAMs that are similar to each other because each VAM would mainly reflect random error. If the differences in the two original VAMs contain information beyond noise (i.e., if some teachers are better at teaching SWDs than others), the variation in the randomized DVAM distribution should be smaller than the original DVAM distribution. This finding would provide evidence that some teachers indeed exhibit a relative advantage for different groups of students that extends beyond random variation. In the context of this paper, this would indicate that there exist individual teachers who are “better” at instructing SWDs compared to non-SWDs.

### **3.4.2 RQ1 Results**

Figure 3.1 provides a scatter plot of each teacher observation by the SWD and non-SWD VAM percentiles. If all teachers were exactly within the same percentile for both student groups (e.g., had the same value-added for each group and no estimation error), the scatter plot would produce a 45-degree line. However, while there is a positive relationship between the percentile of VAMs as seen by the fitted lines, this figure shows that teachers ranked highly in one VAM do not necessarily rank highly in the other. This figure illustrates that some teachers have a relative advantage for one group (i.e., SWD or non-SWD) over another.

**Figure 3.1** Comparison of Teacher SWD and non-SWD VAMs



NOTE: Observations at the teacher level. Math best fit line slope = 0.58. ELA best fit line slope = 0.42. Teacher VAMs shown are averaged across years. VAM (value-added measure). SWD (student with disability).

Table 3.3 further describes the distribution of VAMs across each group by quintile and subject. One finding is that teachers with the highest (or lowest) VAM for one student group tend to also have the highest (or lowest) VAM for the other student group. For example, Panel A of Table 3 shows that the plurality (9.8%) of math teachers fall into the top quintile (i.e., quintile 5) for SWD VAMs and the top quintile for non-SWD VAMs, with the next largest group (9.0%) of teachers falling into

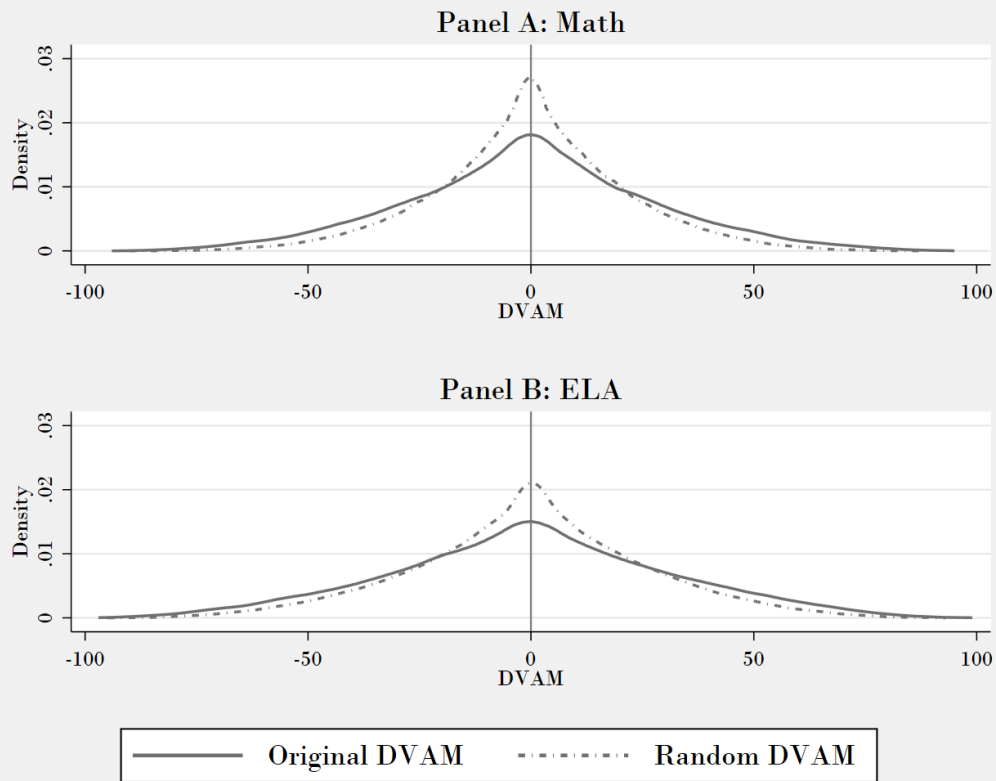
the bottom quintile (i.e., quintile 1) for both VAMs. This pattern persists for ELA teachers, though to a lesser extent. However, for teachers within the middle quintiles, a large majority fall into a bin off the diagonal (i.e., the values in bold) across both subjects. This indicates that teachers closer to the mean for one VAM are more likely to exhibit a relative advantage for SWDs than teachers who fall on the extreme ends of the VAM distribution.

**Table 3.3** Teacher Distribution of SWD and Non-SWD VAMs by Quintile

Panel A: Math		Non-SWD VAM Quintiles				
		1	2	3	4	5
SWD VAM Quintiles	1	<b>9.0%</b>	5.2%	3.2%	1.7%	0.9%
	2	5.0%	<b>5.5%</b>	4.6%	3.2%	1.7%
	3	3.3%	4.5%	<b>4.9%</b>	4.4%	2.9%
	4	2.0%	3.3%	4.5%	<b>5.6%</b>	4.7%
	5	0.7%	1.5%	2.9%	5.1%	<b>9.8%</b>
Panel B: ELA		Non-SWD VAM Quintiles				
		1	2	3	4	5
SWD VAM Quintiles	1	<b>7.4%</b>	5.0%	3.8%	2.6%	1.3%
	2	5.0%	<b>4.8%</b>	4.3%	3.6%	2.3%
	3	3.6%	4.6%	<b>4.4%</b>	4.0%	3.4%
	4	2.6%	3.4%	4.2%	<b>5.0%</b>	4.7%
	5	1.4%	2.3%	3.3%	4.9%	<b>8.2%</b>
Each teacher year observation is assigned to a quintile by its relative rank in the VAM distribution. The lowest quintile of VAMs is represented by 1, and the highest quintile of VAMs is represented by 5. Observations are at the teacher-year level. 34,373 total math observations. 35,197 total ELA observations. SWD (student with disability). VAM (value-added measure). ELA (English language arts).						

In Figure 3.2 we illustrate the relationship between teachers' SWD and non-SWD VAMs for math and ELA courses by plotting the difference between the two VAMs (DVAM). The values further to the left of zero on the x-axis indicate a greater magnitude of a teacher's relative advantage towards SWDs. As previously explained, teachers with a relative advantage for SWDs have a negative DVAM, which indicates that the teacher falls into a higher percentile on the SWD VAM distribution than for the non-SWD VAM.

**Figure 3.2 DVAM and Random DVAM Distributions**



NOTE: This figure compares the DVAM distribution of our sample to the DVAM distribution generated when randomly assigning student with disability (SWD) status within each classroom. DVAM represents the difference in a teacher's percentile ranking in SWD VAM from their percentile ranking in non-SWD VAM. A negative DVAM indicates a relative advantage for instructing SWDs. Since the randomization process preserves the number of SWDs and non-SWDs within each classroom, the *Random DVAMs* use an identical sample as their *Original DVAM* analogue.

To determine the extent to which the original DVAMs are driven by measurement error, we also compare them to a randomized DVAM, represented by the dotted densities in Figure 2. For both subjects, the variation in the original DVAM is greater than the randomized DVAM, which indicates the relative advantage measures for teachers are larger than what we would expect by chance. We compare the variances of the original and random DVAMs to test the differences of the two distributions. The ratios in the variances are 1.70 for math and 1.47 for ELA. In other words, the variance in the original DVAM distribution for math (ELA) courses is 70% (47%) greater than the variance for the random DVAM distribution for the same subject. These results suggest that

the separate SWD and non-SWD VAMs include useful information beyond measurement error and that teachers vary in effectiveness with these two groups of students.

### **3.5 Research Question 2: Are higher SWD/non-SWD VAMs associated with observable teacher characteristics? If so, do teachers with higher SWD VAMs share the same observable characteristics as teachers with higher non-SWD VAMs?**

#### **3.5.1 RQ2 Methods**

The current literature suggests that most observable teacher characteristics (e.g., Master’s degree, certification status) do not accurately predict teacher effectiveness as measured by VAM scores (e.g., Chingos and Peterson, 2011; Kane and Staiger, 2008). However, it may be that observable characteristics are relatively uncorrelated with teacher VAM scores because current VAMs are calculated across all students, rather than by student subgroups. We explore the possibility that observable characteristics of teachers can help identify those who are better at teaching SWDs (or non-SWDs). To do so, we use a standard linear regression model to estimate if observable teacher characteristics are associated with higher/lower VAMs for each subject (math, ELA) and student group (SWD, non-SWD). Our model is as follows:

$$VAM_{jt}^{subject,group} = \beta_0 + \mathbf{X}_{jt}\mathbf{B}_x + \mathbf{Z}_{jt}\mathbf{B}_z + \theta_s + \theta_t + \epsilon_{jt}, \quad (3.3)$$

where  $VAM_{jt}$  is a standardized score calculated using Equation 3.1 for each subject and group.  $\mathbf{X}_{jt}$  is a vector of teacher characteristics, which includes special education credential, Master’s degree or PhD, years of experience, gender, and race.  $\mathbf{Z}_{jt}$  is a vector of classroom characteristics and controls for the share of non-White/non-Asian, English Learner, free or reduced-price lunch eligible, and SWDs within each classroom. We also control for school ( $\theta_s$ ) and year ( $\theta_t$ ) fixed effects to account for any potential differences across schools and yearly shocks, respectively. We estimate a separate regression for each group’s VAM to find the association between VAMs and teacher characteristics. Though the results from Equation 3.3 are purely descriptive, they can

provide valuable insight in determining if higher SWD VAM teachers can be distinguished from higher non-SWD VAM teachers through traditional observable characteristics.

### 3.5.2 RQ2 Results

In general, similar to the existing literature for overall VAMs, we do not find strong discernible relationships between teachers' characteristics and SWD or non-SWD VAMs. Table 3.4 shows the regression results for Equation 3.3 with each column displaying the coefficients for teacher characteristics across each subject-group combination of VAM. While the special education credential status, education, and experience of teachers are largely uncorrelated with VAMs, there are some differences by teacher gender and race. Across both subjects, female teachers are associated with higher non-SWD VAMs, while there does not seem to be any relationship with SWD VAMs. For teacher race, relative to white teachers, on average, Asian and Latino teachers are associated with higher math VAMs for both student groups, while Black teachers are associated with lower math VAMs. These differences in VAMs between teachers of color and their White counterparts are larger for non-SWD VAMs than for SWD VAMs. For example, while Asian teachers exhibit higher math SWD VAMs than White teachers (by 0.140 SD), Asian teachers are expected to have even higher math non-SWD VAMs (by 0.170 SD). The differences across VAM types are statistically significant for Asian and Black teachers, but they are not for Latino teachers.<sup>5</sup>

In general, we do not find any statistical differences for ELA VAMs by teacher race. The one exception is that, on average, Black teachers exhibit lower non-SWD ELA VAMs than White teachers. While it is not clear why these relationships emerge, one potential explanation is that, given LAUSD's large Latino population (as seen in Table 3.1, Latino students represent 73.8% of the sample), a student-teacher race match effect could be influencing the VAMs (e.g., Dee, 2004; Egalite et al., 2015; Joshi et al., 2018; Wood and Lai, 2022). While this is an intriguing hypothesis, a deeper analysis of this finding is beyond the scope of this study and so we leave it for future

---

<sup>5</sup>We use a Wald test to determine if any differences found between the coefficients for each teacher race across VAM types are statistically significant. P-values for the differences in the math SWD VAM and non-SWD VAM coefficients are 0.0027, 0.0000, and 0.1856 for Asian, Black, and Latino/a, respectively.

**Table 3.4** Association between Teacher Characteristics and SWD and non-SWD VAMs

	Math VAM		ELA VAM	
	(1) SWD	(2) Non-SWD	(3) SWD	(4) Non-SWD
SpEd Credential	-0.182 (0.094)	-0.188 (0.107)	-0.002 (0.094)	-0.068 (0.071)
Master's or PhD	-0.009 (0.026)	-0.024 (0.025)	0.062* (0.025)	0.022 (0.024)
Exp: Early Career (3-5 yrs)	0.002 (0.038)	0.109** (0.040)	0.033 (0.038)	0.032 (0.043)
Exp: Middle Career (6-9 yrs)	0.015 (0.045)	0.074 (0.042)	0.026 (0.044)	0.007 (0.044)
Exp: Late Career (10+ yrs)	-0.131** (0.048)	-0.063 (0.044)	-0.015 (0.042)	-0.015 (0.043)
Female	0.031 (0.031)	0.097** (0.030)	0.018 (0.029)	0.211*** (0.030)
Asian	0.140** (0.047)	0.170*** (0.045)	-0.072 (0.047)	0.060 (0.045)
Black	-0.101* (0.048)	-0.271*** (0.050)	-0.054 (0.040)	-0.253*** (0.042)
Latino	0.077* (0.034)	0.068* (0.033)	0.032 (0.030)	0.031 (0.031)
Other Race	0.045 (0.070)	-0.015 (0.065)	-0.056 (0.068)	-0.035 (0.073)
N	34,373	34,373	35,197	35,197

School-clustered standard errors shown in parentheses. All columns include school and year fixed effects and classroom demographic characteristics. Novice teachers (1-2 yrs) are excluded experience group. White teachers are excluded racial/ethnic group. SWD (student with disability). VAM (value-added measure). ELA (English language arts). SpEd (special education). Exp (years of experience). \*p<.05. \*\*p<.01. \*\*\*p<.001.

research.

### 3.6 Research Question 3: Do SWDs sort into classes taught by a teacher with high (or low) SWD VAMs?

#### 3.6.1 RQ3 Methods

To investigate whether SWDs sort into classrooms taught by teachers with high (or low) SWD VAMs, for all students, we regress their teacher's VAM on SWD status using the following equation:

$$Y_{it}^{subject,group} = \alpha_0 + \alpha_1 SWD_{it} + \theta_{sgt} + \epsilon_{it}, \quad (3.4)$$

where  $Y$  represents a teacher's VAM for a given subject (math, ELA) and group of students (SWD, non-SWD) for a given year  $t$ . In this equation,  $SWD$  indicates if student  $i$  is a SWD in year  $t$ .  $\theta_{sgt}$  is a school-grade-year fixed effect and allows us to focus specifically on how SWDs and non-SWDs are being sorted within individual school-grade-year combinations, since this is the level at which student assignment to teachers occurs. The coefficient of interest,  $\alpha_1$ , is interpreted as the expected difference in teacher VAM for SWDs compared to non-SWDs. We conduct two additional analyses. First, we check for heterogeneity in sorting across disability types by replacing the SWD indicator variable in Equation 3.4 with indicators for a specific disability type (e.g., autism, specific learning disability, speech or language impairment, or other disability).

$$Y_{it}^{subject,group} = \alpha_0 + \alpha_1 SWD_{it} X_{it} + \alpha_2 SWD_{it} + \alpha_3 X_{it} + \theta_{sgt} + \epsilon_{it}, \quad (3.5)$$

Additionally, in Equation 3.5, we interact the SWD indicator with different student characteristics to further explore potential heterogeneity by FRL eligibility, EL status, or non-White/Asian status.  $X$  represents these student characteristics.



### 3.6.2 RQ3 Results

Table 3.5 provides results from Equation 3.4 where we estimate the expected difference in teacher VAMs for SWDs compared to non-SWDs. For both subjects, we show the results for SWD VAMs and non-SWD VAMs. We find that SWDs are consistently assigned to teachers with lower VAMs. Note that these are teacher-level VAMs and so they imply that a SWD is placed with a teacher who ranks 0.036 and 0.043 standard deviations (SD) lower in the teacher distribution of SWD and non-SWD value-added, respectively. The analogous estimates for ELA VAMs are .040 and .056 SD, reflecting a similar pattern where SWDs sort into classrooms taught by teachers with lower value-added. All estimates are statistically significant at the 0.1% level.<sup>6</sup>

**Table 3.5** Association between Student Disability Status and Teacher VAMs

	Math VAM		ELA VAM	
	(1) SWD	(2) Non-SWD	(3) SWD	(4) Non-SWD
SWD	-0.036*** (0.005)	-0.043*** (0.004)	-0.040*** (0.006)	-0.056*** (0.004)
N	1,408,173	1,408,173	1,481,418	1,481,418
Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year fixed effects. SWD (student with disability). VAM (value-added measure). *p<.05. **p<.01. ***p<.001.				

We conduct a similar analysis to Table 3.5 but further disaggregate by type (autism, SLD, LS Imp., and other disabilities) in Table H.1 of Appendix H. For both subjects, each column represents a separate regression where SWD VAM and non-SWD VAM are outcomes. For students with SLD and other disabilities, a pattern similar to Table 5 holds: on average, these students are assigned to teachers with lower VAMs than their non-SWD peers. Relative to coefficients for all SWDs, the coefficients are slightly larger in magnitude for SLD students and slightly smaller for students with other disabilities. On the other hand, students with autism or LS Imp. exhibit little difference in

<sup>6</sup>We conduct a similar analysis to Table 4 using three different samples in Table G.1 of Appendix G. In this analysis, we vary the minimum number of student observations that a teacher must have for each student group (SWD, non-SWD) for a subject. We find similar results across all samples, so henceforth, we only show results using a 10-student minimum.

teacher value-added relative to their non-SWD peers.

Table 3.6 provides results from Equation 3.5, which furthers the analysis above by examining heterogeneity across different student characteristics. We analyze FRL eligibility in Panel A, EL status in Panel B, and non-White/non-Asian status in Panel C. Across all panels, we find that SWDs that are eligible for FRL, are ELs, or are non-White/non-Asian are assigned to teachers with lower VAMs than non-SWDs without these characteristics (i.e., non-FRL, non-EL, White or Asian non-SWDs). For instance, on average, a non-White/non-Asian SWD has a math teacher with a SWD VAM that is .087 SD lower than a White or Asian non-SWD. Moreover, our estimates suggest that SWDs that are also either (1) FRL-eligible or (2) non-White/non-Asian have teachers with lower VAMs than students that have just one of these characteristics (e.g., non-FRL SWDs, or non-White/non-Asian non-SWDs). This finding suggests that for some groups of SWDs, there may be a compounding pattern of disadvantage, although this effect is relatively small.

### **3.7 Research Question 4: Are SWDs sorted to teachers with relatively higher SWD VAMs compared to their non-SWD VAMs?**

#### **3.7.1 RQ4 Methods**

The previous research question examines whether SWDs sort into classrooms taught by teachers with higher SWD VAMs. Our fourth research question focuses on the DVAM measure to understand how SWDs sort across teachers by teacher relative advantage. In other words, we ask whether SWDs sort towards teachers with higher SWD VAMs relative to their non-SWD VAMs, even if a teacher has low VAMs for both student groups. To answer this question, we indicate a teacher as having a relative advantage for SWDs using the DVAM measure described in *Research Question 1* as the outcome to Equation 3.4. If a teacher's DVAM is  $<0$ , they have a relative advantage in teaching SWDs; otherwise, they do not. As a robustness check to our definition of relative advantage, we also create two additional measures where teachers must have more than a 10 or 20 percentile advantage in teaching SWDs (i.e., a DVAM of  $<-10$  and  $<-20$ ) to be defined as having a relative advantage. In these two additional definitions of relative advantage, a teacher with a DVAM of  $-5$

**Table 3.6** Association between Student Disability Status and Teacher VAMs

Panel A: FRL	Math VAM		ELA VAM	
	SWD	Non-SWD	SWD	Non-SWD
FRL	-0.034*** (0.007)	-0.040*** (0.005)	-0.036*** (0.007)	-0.048*** (0.005)
SWD	-0.041*** (0.010)	-0.053*** (0.008)	-0.059*** (0.010)	-0.081*** (0.007)
FRL $\times$ SWD	0.008 (0.009)	0.013 (0.007)	0.024* (0.009)	0.032*** (0.007)
<i>Sum of coeffs.</i>	-0.067*** (0.010)	-0.080*** (0.007)	-0.071*** (0.010)	-0.097*** (0.007)
Panel B: EL	Math VAM		ELA VAM	
	SWD	Non-SWD	SWD	Non-SWD
EL	-0.047*** (0.009)	-0.058*** (0.007)	-0.020 (0.011)	-0.062*** (0.008)
SWD	-0.041*** (0.006)	-0.051*** (0.005)	-0.044*** (0.007)	-0.068*** (0.005)
EL $\times$ SWD	0.037*** (0.007)	0.050*** (0.006)	0.019* (0.008)	0.061*** (0.006)
<i>Sum of coeffs.</i>	-0.051*** (0.008)	-0.059*** (0.007)	-0.045*** (0.011)	-0.069*** (0.008)
Panel C: Non-White, non-Asian	Math VAM		ELA VAM	
	SWD	Non-SWD	SWD	Non-SWD
Non-White, non-Asian	-0.057*** (0.010)	-0.071*** (0.008)	-0.060*** (0.010)	-0.085*** (0.008)
SWD	-0.061*** (0.013)	-0.067*** (0.010)	-0.065*** (0.012)	-0.091*** (0.008)
Non-white, non-Asian $\times$ SWD	0.030* (0.012)	0.029** (0.009)	0.029* (0.011)	0.042*** (0.008)
<i>Sum of coeffs.</i>	-0.087*** (0.012)	-0.109*** (0.009)	-0.096*** (0.013)	-0.134*** (0.010)
N	1,408,173	1,408,173	1,481,418	1,481,418

Teacher-clustered standard errors in parentheses. All columns include school-grade-year fixed effects. *Sum of coeffs.* show difference in VAM for SWD with the specific characteristic and non-SWD without characteristic. SWD (student with disability). VAM (value-added measure). FRL (free or reduced-price lunch). EL (English learner). \*p<.05. \*\*p<.01. \*\*\*p<.001.

would not be said to have a relative advantage for SWDs. However, a teacher with a DVAM of -25 would be defined as having a relative advantage for SWDs. As in *Research Question 3*, we repeat the analysis using indicators for each disability type.

### **3.7.2 RQ4 Results**

Panel A of Table 3.7 provides estimates from Equation 3.4 where the outcome is an indicator variable for relative advantage in teaching SWDs. For math courses, we do not find evidence that SWDs are more likely to be placed with teachers who have a relative advantage in teaching SWDs compared to non-SWDs. However, the results are stronger for ELA; using any of our three definitions of relative advantage, SWDs are 0.6-0.8 percentage points more likely to be assigned to teachers with a relative advantage in teaching SWDs than non-SWDs. When viewed in the context of our results from Table 3.5, we find that on average, while SWDs have teachers with lower VAMs, SWDs are more likely to be placed with ELA teachers who are comparatively more effective in teaching SWDs.

We perform an analysis similar to that in Panel A of Table 3.7 that focuses on differences across SWD disability type, shown in Table H.2 of Appendix H. For ELA, the results are relatively similar to those in Table 3.7 for students with SLD and other disabilities. The results are not statistically significant for students with autism or LS Imp. For math, while we find a positive effect for students with SLD and other disabilities using our least restrictive definition of relative advantage, this impact is statistically insignificant when using our most strict definition. In general, we do not find statistically significant results for students with autism or LS Imp.

The relative advantage measure in Panel A of Table 3.7 is defined across school-grades for the entire district. In Panel B, the outcome is a dummy variable which indicates teachers whose DVAMs are strictly less than (i.e., have more relative advantage for instructing SWDs) the median DVAM of all teachers within a given school-grade-year combination. In this approach, we assess the relationship between SWDs and the relative advantage measure while only considering teachers with which the student could feasibly share a classroom. Looking within schools strengthens

**Table 3.7** Association between Student Disability Status and Teacher Relative Advantage

Panel A: SWD sorting to teachers with relative advantage toward instructing SWDs						
Relative Advantage Definition	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
	>0	>10	>20	>0	>10	>20
SWD	0.004 (0.002)	0.004 (0.002)	0.003 (0.002)	0.007** (0.003)	0.008** (0.003)	0.006** (0.002)
N	1,408,173	1,408,173	1,408,173	1,481,418	1,481,418	1,481,418
Panel B: SWD sorting to teachers below median DVAM <i>within</i> school-grade-year						
	Math			ELA		
	(1)			(2)		
SWD	0.007** (0.002)			0.010*** (0.003)		
N	1,408,173			1,481,418		
In Panel A, the outcome variable is a dummy indicator for relative advantage. >0 (>10; >20) indicate that the SWD VAM percentile is more than 0 (10; 20) percentile points greater than the non-SWD VAM percentile. In Panel B, the outcome is a dummy indicator for teachers whose DVAMs are strictly less than the median DVAM of all teachers within a given school-grade-year combination. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year fixed effects. SWD (student with disability). ELA (English language arts). VAM (value-added measure). *p<.05. **p<.01. ***p<.001.						

our previous findings. Across both subjects, SWDs are more likely than non-SWDs (by 0.7 and 1.0 percentage points for math and ELA, respectively) to sort into classrooms taught by teachers whose relative advantage is more geared towards SWDs than the other teachers in the school-grade. These results suggest that from a school's perspective, there may be room for efficiency gains from assigning SWDs to teachers who are comparatively stronger at teaching SWDs. We note, however, that from a student or parent's perspective, any notion of efficiency may be outweighed by overall teacher efficacy, since SWDs still receive lower VAM teachers on average.

### 3.8 Research Question 5: How do retention rates compare for teachers with high versus low SWD VAMs? Are schools more or less likely to retain teachers with relative advantages for instructing SWDs than those with relative disadvantages?

#### 3.8.1 RQ5 Methods

Whether or not a student is assigned to a highly effective teacher (as measured by a teacher's VAM score for each subject and student group) is, to some extent, determined by whether a school is able to retain such teachers. We leverage indicator variables for whether a teacher leaves LAUSD or switches schools within the district.<sup>7</sup> Using Equation 3.6, we measure how an increase in SWD and non-SWD VAMs relate to a teacher's probability of leaving the district or switching schools. Specifically, we use the following model to understand the association between teacher mobility and each group-VAM for both subjects:

$$Y_{jt} = \alpha_0 + \alpha_1 VAM_{jt}^{subject, SWD} + \alpha_2 VAM_{jt}^{subject, non-SWD} + \mathbf{X}_{jt} \mathbf{B}_x + \mathbf{Z}_{jt} \mathbf{B}_z + \theta_s + \theta_t + \epsilon_{jt}, \quad (3.6)$$

where  $Y_{jt}$  indicates one of the two outcomes (i.e., leave LAUSD or switch schools). The remaining components are defined as in previous models.

Additionally, we use the following model to analyze whether teachers with a relative advantage in teaching SWDs are more or less likely to leave LAUSD or switch schools than other teachers:

$$Y_{jt} = \alpha_0 + \alpha_1 RelAdv_{jt}^{subject} + \mathbf{X}_{jt} \mathbf{B}_x + \mathbf{Z}_{jt} \mathbf{B}_z + \theta_s + \theta_t + \epsilon_{jt}, \quad (3.7)$$

where  $RelAdv_{jt}^{sub}$  is an indicator variable representing a teacher with a relative advantage for teaching SWDs. Like the analysis from *Research Question 4*, we check the robustness of our results across different definitions of relative advantage for instructing SWDs.

---

<sup>7</sup>Teachers that leave the district due to reduction in force (layoffs) are excluded from the leave LAUSD analysis. Also, we exclude teachers that leave the district from the switch school analysis.

### 3.8.2 RQ5 Results

Panel A of Table 3.8 describes the relationship between the probability that a teacher leaves the district and teacher VAMs. The first column shows the regression results for Equation 3.6 but only includes SWD math VAMs. We find that teachers with higher SWD math VAMs leave the school district at a lower rate than average. Increasing a teacher's SWD math VAM by one SD is associated with a lower probability of leaving LAUSD by 0.20 percentage points. Since the average probability of any teacher leaving the district is 2.06 percent, this finding represents about 9.7 percent of the mean. The second column analyzes the relationship between non-SWD math VAMs and teachers leaving LAUSD, and we find similar results to the first column. That is, for each one SD increase in a teacher's non-SWD math VAM, the teacher's likelihood of leaving the district decreases by 0.26 percentage points on average. The third column analyzes Equation 3.6 which includes both student group VAMs. While the non-SWD math VAM result remains similar, the finding for SWD math VAMs is statistically insignificant. For ELA, we find a negative association between higher non-SWD VAMs and teachers leaving the district, but do not find any relationship for SWD VAMs. Overall, these results indicate that LAUSD retains teachers effective at instructing non-SWDs. We repeat this analysis using an indicator for whether a teacher switches schools within the district in Table I.1 in Appendix I. In general, we do not find much evidence of a relationship between VAMs and switching schools. The lone exception is that we find that teachers with higher non-SWD math VAMs tend to switch schools at lower rates than average teachers.

In Panel B of Table 3.8, we analyze Equation 3.7 which uses a measure of relative advantage as the regressor. We do not find any strong relationships between relative advantage and the likelihood of teachers leaving the district. While the effect found in column (4) suggests a significant relationship between relative ELA advantage and increased likelihood of leaving LAUSD, this finding is not robust to more strict definitions of relative advantage. Panel B of Table I.1 presents the relationships between relative advantage and teachers' likelihood of switching schools. Our estimates suggest that math teachers with a relative advantage (defined either with the 10- or 20-percentile threshold) in teaching SWDs are 0.69 percentage points more likely to switch schools

**Table 3.8** VAMs and Relative Advantage Association with Leaving LAUSD

Panel A: VAM	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
SWD VAM	-0.0020** (0.0006)		-0.0008 (0.0007)	-0.0005 (0.0006)		0.0003 (0.0007)
Non-SWD VAM		-0.0026*** (0.0006)	-0.0022** (0.0007)		-0.0020** (0.0006)	-0.0021** (0.0007)
Panel B: Relative Advantage	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Advantage (RA)	0.0025 (0.0014)	0.0012 (0.0016)	0.0007 (0.0018)	0.0031* (0.0014)	0.0026 (0.0016)	0.0028 (0.0017)
<i>RA Definition</i>	>0	>10	>20	>0	>10	>20
N	34,355	34,355	34,355	35,154	35,154	35,154

Leave LAUSD is a dummy indicator that equals 1 when a teacher leaves Los Angeles Unified School District at the end of the academic year. Relative advantage variables are dummy indicators that equal 1 when a teacher has a SWD VAM percentile > non-SWD VAM percentile for a given year in that subject. >0 (>10; >20) indicate that the SWD VAM percentile is more than 0 (10; 20) percentile points greater than the non-SWD VAM percentile. Teachers that leave due to reduction in force are excluded. Teacher-clustered standard errors shown in parentheses. All columns include school and year fixed effects. Coefficients for teacher and class characteristics not shown. VAM (value-added measure). SWD (student with disability). ELA (English language arts). RA (relative advantage). \*p<.05. \*\*p<.01. \*\*\*p<.001.

than their peers, while ELA teachers with a relative advantage do not have a strong relationship with switching schools.

### 3.9 Discussion

In this study, we seek to understand whether teachers are differentially effective at producing academic gains for students with and without disabilities. We also explore whether SWDs are more likely to be assigned to teachers who are more effective at supporting them and whether teachers who are differentially effective for SWDs stay in LAUSD. The question of whether a teacher is differentially effective with one student subgroup over another seems relatively straightforward, but we know of no previous studies that have provided empirical data to inform this question with respect to students with disabilities. One challenge is having a sufficiently large longitudinal sample in order to calculate – with reasonable precision – estimates of teachers’ effectiveness within each group. We address this by using data from the Los Angeles Unified School District from 2007-



2008 to 2017-2018. A second challenge in examining differential effectiveness is distinguishing between signal and noise. When calculating two different VAMs for the same teacher, it is hard to know whether the results are attributable to differences in effectiveness or whether they reflect measurement error. To address this concern, we calculated differences in VAMs based on randomly assigned SWDs and non-SWDs, giving us some sense of whether differences in effectiveness that we see in our data are wider or narrower than those we would expect to see by chance.

We have several key findings. First, we establish that teachers are differentially effective at supporting SWDs. Teachers who are effective at teaching non-SWDs are not synonymous with teachers who are best able to support SWDs. Indeed, we show that a sizeable share of the top performing teachers for non-SWDs have relatively poor performance for SWDs (as measured by VAMs). Second, we started with the assumption that it may be preferable for principals to identify which teachers have an advantage in teaching specific student groups when making classroom assignments. Hence, we consider whether our VAMs for each group are related to observable teacher characteristics. Unfortunately, there does not seem to be any relationship between either SWD or non-SWD VAMs and teacher education, experience, or credentials. Third, we find that SWDs have teachers with lower VAMs on average. These findings are consistent across both subjects and VAM types. Fourth, we turn to considering how students sort to teachers based on relative advantage. If some teachers are particularly effective with SWDs, we would ideally see that there is some sorting of SWDs into their classrooms. While we find that SWDs tend to sort to lower performing teachers overall, relative to non-SWDs, we find that SWDs are slightly more likely to have a teacher who is more effective at teaching SWDs than non-SWDs in ELA courses – e.g., she has a relative advantage teaching SWDs. In other words, for ELA courses, SWDs are likely to have lower VAM teachers, but these teachers do appear to be more effective with their SWDs than their non-SWDs. This pattern strengthens to include both subjects when we analyze the relative advantage measure within a school. This suggests that SWDs in LAUSD sort in a way that, while having lower performing teachers, nonetheless exhibits some efficiency benefits.

Overall, the results suggest that SWDs end up with relatively lower performing teachers regard-

less of how we measure value-added. Consistent with existing literature (e.g., Goldhaber et al., 2011; Hanushek et al., 2016), high non-SWD VAM teachers are less likely to leave LAUSD than low VAM teachers. However, we do not find evidence of a relationship between SWD VAM and teachers leaving the district. Thus, only the most effective teachers for students without disabilities are more likely to be retained. The remaining challenges are retaining teachers who are effective at teaching SWDs and ensuring that students with disabilities gain access to those teachers.

Collectively, how should we think about these results? Schools face a legal mandate to ensure that they are promoting equitable outcomes for students with disabilities (*Endrew F. vs. Douglas County School District RE-1*, 2017). It is essential that schools have the information necessary to ensure that SWDs are being assigned to teachers most likely to promote positive academic achievement. As our study demonstrates, overall VAMs likely do not provide enough information to inform this question. SWD-specific VAMs may be more suitable for identifying which teachers have a history of effectiveness with SWDs and whether students are being optimally assigned to these teachers. We see some evidence that SWDs are being assigned to teachers who are differentially effective at supporting this student subgroup. One silver lining is that these decisions have been made absent of the kinds of SWD-specific information we include in this study. However, there remains room for improvement. Moving forward, we encourage schools, districts, and policymakers to consider how data systems might be organized to inform a more systematic and efficient assignment of SWDs to teachers who are most likely to promote equitable academic outcomes.

## **APPENDICES**

## APPENDIX A

### STUDENT DESCRIPTIVE CHARACTERISTICS AND NONCOGNITIVE SKILL CORRELATIONS

**Table A.1** Average of Teacher Characteristics by Student Race

	Asian	Black	Latino	White	Total
Share Asian Tch.	0.140 (0.158)	0.085 (0.132)	0.089 (0.134)	0.099 (0.140)	0.092 (0.137)
Share Black Tch.	0.070 (0.123)	0.257 (0.265)	0.098 (0.164)	0.063 (0.119)	0.108 (0.177)
Share Latino Tch.	0.196 (0.198)	0.170 (0.187)	0.319 (0.251)	0.174 (0.182)	0.285 (0.245)
Share White Tch.	0.550 (0.247)	0.435 (0.271)	0.435 (0.264)	0.620 (0.236)	0.458 (0.267)
Share Other Tch.	0.044 (0.095)	0.053 (0.106)	0.059 (0.111)	0.045 (0.096)	0.057 (0.109)
Share Female Tch.	0.520 (0.232)	0.525 (0.241)	0.508 (0.237)	0.535 (0.235)	0.513 (0.237)
Share Fully Cert. Tch.	0.928 (0.132)	0.860 (0.200)	0.881 (0.183)	0.928 (0.133)	0.886 (0.179)
Share Novice Tch.	0.031 (0.086)	0.065 (0.134)	0.057 (0.124)	0.028 (0.082)	0.054 (0.121)
Share Early Career Tch.	0.073 (0.128)	0.103 (0.157)	0.102 (0.158)	0.065 (0.123)	0.097 (0.154)
Share Mid Career Tch.	0.104 (0.146)	0.126 (0.163)	0.127 (0.165)	0.105 (0.148)	0.123 (0.162)
Share Late Career Tch.	0.792 (0.215)	0.706 (0.257)	0.714 (0.258)	0.802 (0.213)	0.726 (0.254)
N	112,037	245,796	2,016,692	212,427	2,674,791

Observations at student-year level. The mean of each characteristics is displayed with the standard deviations in parentheses. *Share [Race] Tch.* represents the average share of [race] teachers that a student has in a given year; *Fully Cert.* indicates whether a teacher is fully certified; *Novice* indicates a teacher with less than or equal to 2 years of experience; *Early Career* indicates 3 through 5 years of experience; *Mid Career* indicates 6 through 7 years of experience; *Late Career* indicates more than 8 years of experience.

**Table A.2** Summary of Share Race Match and Student Noncognitive Skills Measures

	Asian	Black	Latino	White	Total
Share RM	0.140 (0.158)	0.257 (0.265)	0.319 (0.251)	0.620 (0.236)	0.321 (0.266)
GPA	3.168 (0.789)	2.210 (0.950)	2.318 (0.968)	2.890 (0.919)	2.408 (0.983)
Work Habits	1.610 (0.433)	1.045 (0.523)	1.150 (0.534)	1.440 (0.518)	1.192 (0.543)
Cooperation	1.743 (0.326)	1.225 (0.481)	1.360 (0.457)	1.613 (0.410)	1.393 (0.464)
Held Back	0.013 (0.115)	0.031 (0.173)	0.041 (0.198)	0.018 (0.131)	0.036 (0.187)
Suspension	0.012 (0.107)	0.102 (0.302)	0.037 (0.189)	0.026 (0.158)	0.040 (0.197)
Days Absent	4.406 (8.121)	10.439 (13.481)	8.379 (12.649)	8.102 (10.422)	8.297 (12.359)
N	112,037	245,796	2,016,692	212,427	2,674,791

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 A.3** Cross-correlations of Noncognitive Components

	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

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.

## APPENDIX B

### CRITERIA FOR MARKS IN LAUSD

**Figure B.1** Rubric for Marks for Achievement, Work Habits, and Cooperation



**LOS ANGELES UNIFIED SCHOOL DISTRICT  
POLICY BULLETIN**

**ATTACHMENT A**

#### CRITERIA FOR MARKS

Academic Mark	A	B	C	D	FAIL
<b>Quality of Work</b>	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.
<b>Interpretation and Application</b>	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.
<b>Thinking and Reasoning Skills</b>	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.
<b>Quantity of Work</b>	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
<b>Effort</b>	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.
<b>Responsibility</b>	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.
<b>Attendance</b>	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.
<b>Evaluation</b>	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
<b>Courtesy</b>	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.
<b>Conduct</b>	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.
<b>Improvement</b>	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.
<b>Class Relations</b>	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.

## APPENDIX C

### EXPLORATORY FACTOR ANALYSIS

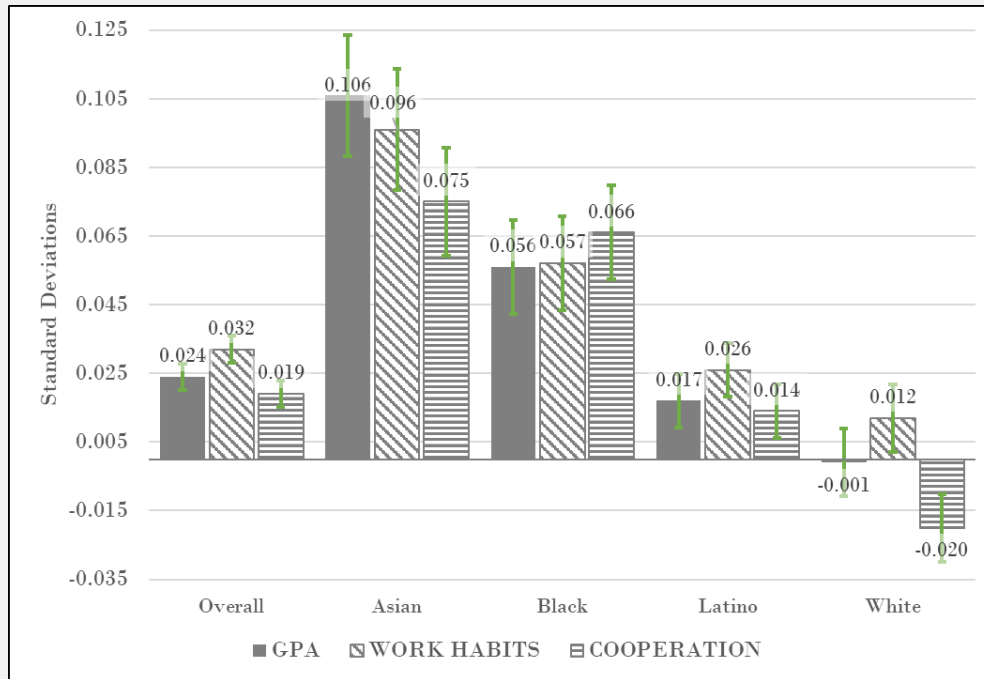
**Table C.1** Factor Loadings and Eigenvalues for Exploratory Factor Analysis

	Noncog. Index	Cognitive	Uniqueness
GPA	0.90	0.28	0.11
Work Habits	0.98	0.18	0.00
Cooperation	0.86	0.24	0.21
Held Back	-0.10	-0.06	0.99
% Days Absent	-0.37	-0.13	0.84
Suspension	-0.20	-0.07	0.95
Math Test	0.37	0.75	0.29
ELA Test	0.37	0.76	0.29
Eigenvalue	2.97	1.34	
Sample includes students from Los Angeles Unified School District (LAUSD) in grades 6 through 12 for academic years 2007-08 through 2017-18 with math and ELA test scores. Observations are at the student-year level. Figure shows the component loadings post varimax rotation across the noncognitive index (eigenvalue = 2.97) and the cognitive index (eigenvalue = 1.34).			

## APPENDIX D

### RACE MATCHED EFFECTS FOR NONCOGNITIVE COMPONENTS

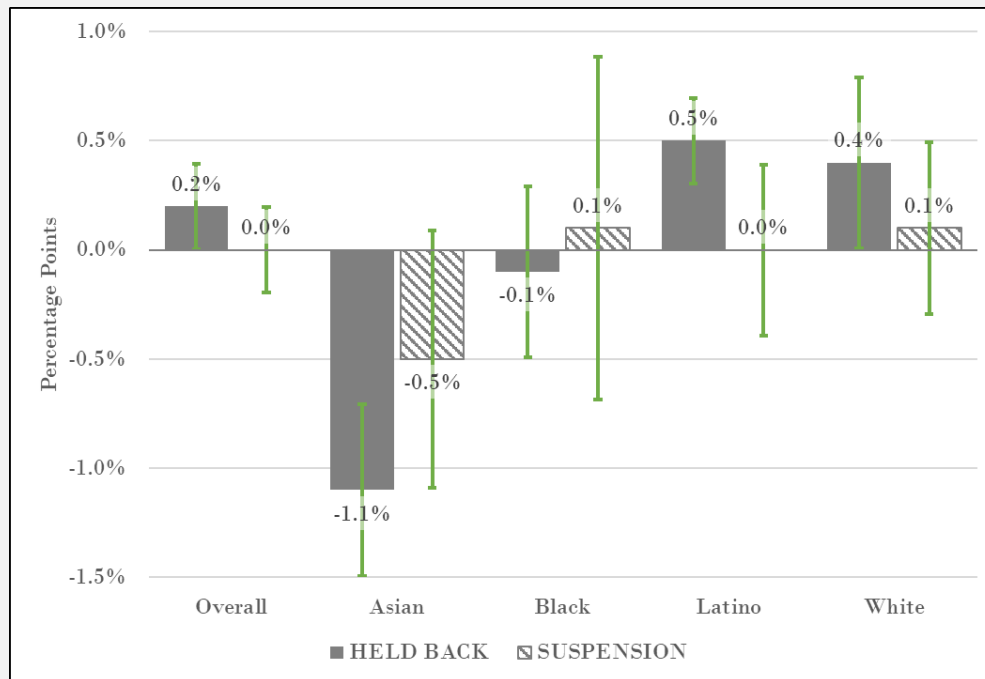
**Figure D.1** Race Match Effect on Standardized Noncognitive Measures



Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered, and the whiskers show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* shows the results for Equation 1.1 and was analyzed separately from the remaining groups. Equation 1.2 was used to analyze the race match effects by student race. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the year level. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure is analogous to Table 1.3 in Section 1.4.

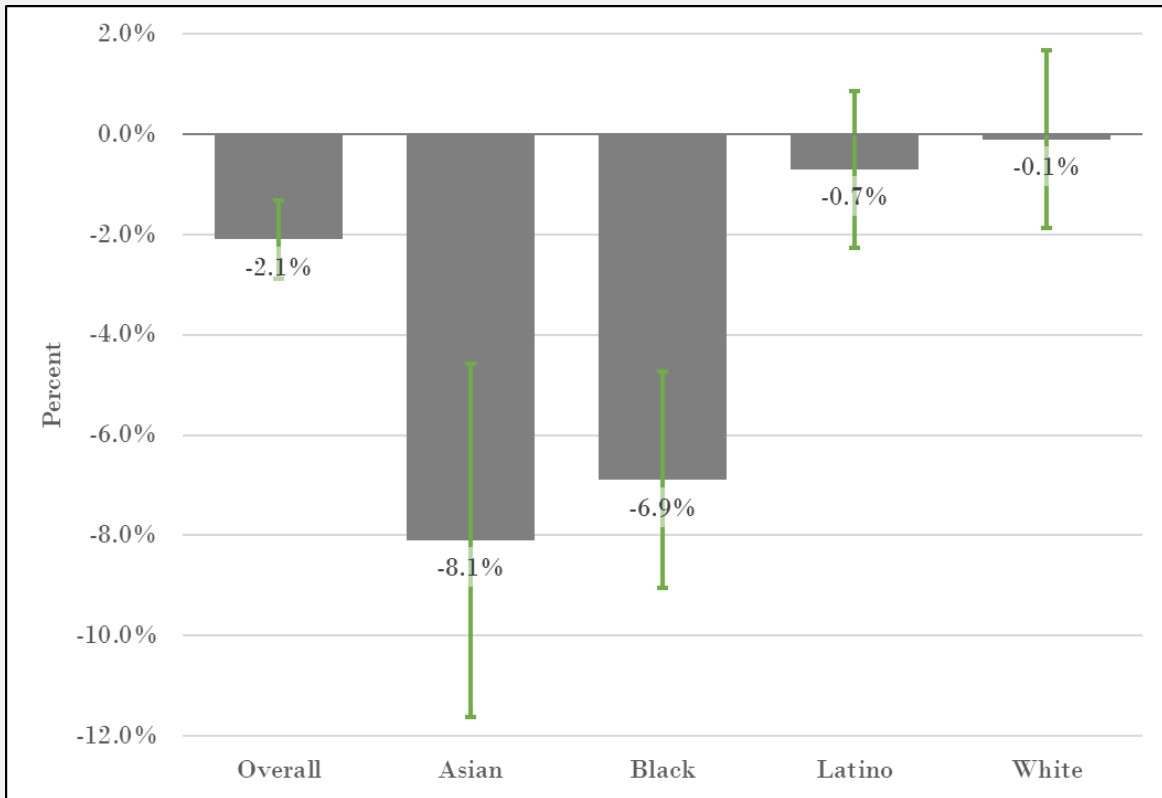


**Figure D.2 Race Match Effect on the Probability of Noncognitive Measures**



Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered, and the whiskers show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* shows the results for Equation 1.1 and was analyzed separately from the remaining groups. Equation 1.2 was used to analyze the race match effects by student race. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *Held Back* and *Suspension* are dummy variables and indicate if a student has been held back or suspended for the given year. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure is analogous to Table 1.3 in Section 1.4.

**Figure D.3 Race Match Effect on  $\ln(\text{Days Absent})$**



Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered, and the whiskers show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* shows the results for Equation 1.1 and was analyzed separately from the remaining groups. Equation 1.2 was used to analyze the race match effects by student race. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics.  $\ln(\text{Days Absent})$  is calculated using  $\ln(x+1)$  to account for zeros. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure is analogous to Table 1.3 in Section 1.4.

## APPENDIX E

### SUMMARY STATISTICS, RESULTS, AND PLACEBO TEST FOR COURSE LEVEL ANALYSIS

**Table E.1** Percent of Students with Direct Race Match by Subject

	Asian	Black	Latino	White
Overall	14.2	25.7	31.7	61.3
ELA	10.5	24.9	35.0	64.6
Math	21.6	27.1	32.3	54.7
History	9.4	23.9	34.6	64.4
Science	18.8	25.0	25.7	58.0
Health	14.0	29.6	33.1	59.0

Observations are at the student-year-course level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Each column shows the percent of directly race matched students within the given race across subjects. The first row shows the overall race match percentages by student race. A *direct race match* refers to students sharing a class with a same race teacher. This study uses *course* and *subject* interchangeably. History contains "social studies" courses, and Health contains "physical education" courses.

**Table E.2** Percent of Teachers by Subject across Teacher Race

	Asian	Black	Latino	White
ELA	5.6	12.7	30.8	47.2
Math	13.0	13.9	28.2	35.9
History	6.3	12.4	29.2	47.6
Science	11.8	13.1	22.9	42.2
Health	8.2	14.9	28.6	43.9

Observations are at the teacher-year-grade-course level for school years 2007-08 through 2017-18 and grades 6 through 12. Each column shows the percent of teacher observations across the given subjects by teacher race. This study uses *course* and *subject* interchangeably. History contains "social studies" courses, and Health contains "physical education" courses.

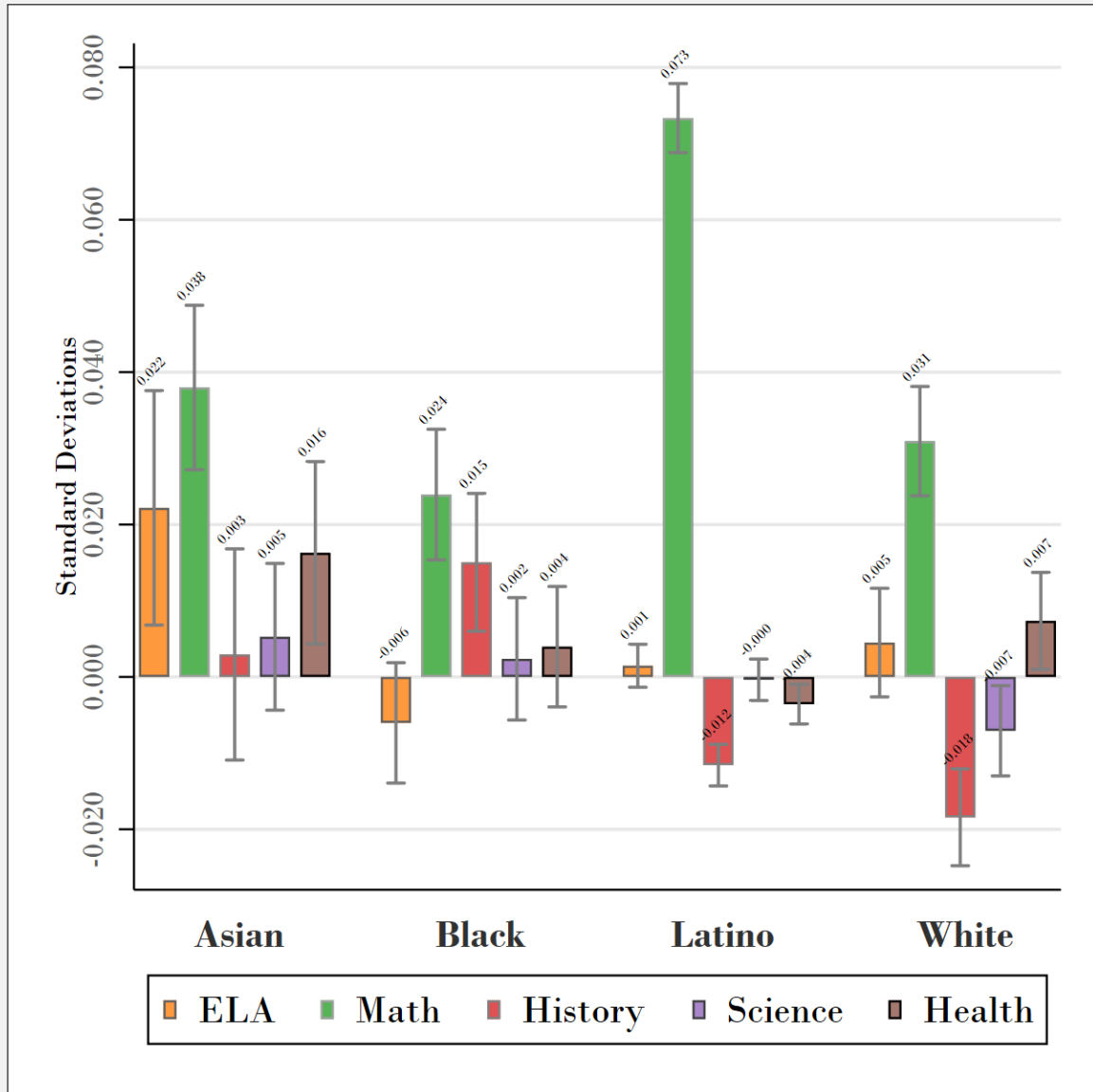
**Table E.3** Direct Race Match Effect at Course Level

	GPA				Work Habits				Cooperation			
	(1) Main	(2) Stu FE	(3) Stu-Sub FE	(4) Stu-Yr FE	(5) Main	(6) Stu FE	(7) Stu-Sub FE	(8) Stu-Yr FE	(9) Main	(10) Stu FE	(11) Stu-Sub FE	(12) Stu-Yr FE
Asian*RM	0.106*** (0.009)	0.090*** (0.003)	0.081*** (0.003)	0.087*** (0.002)	0.096*** (0.009)	0.073*** (0.003)	0.068*** (0.003)	0.069*** (0.003)	0.075*** (0.008)	0.057*** (0.003)	0.049*** (0.003)	0.056*** (0.002)
Black*RM	0.056*** (0.007)	0.063*** (0.002)	0.061*** (0.002)	0.062*** (0.002)	0.057*** (0.007)	0.051*** (0.002)	0.049*** (0.002)	0.049*** (0.002)	0.066*** (0.007)	0.058*** (0.002)	0.056*** (0.002)	0.055*** (0.002)
Latino*RM	0.017*** (0.004)	0.046*** (0.001)	0.049*** (0.001)	0.051*** (0.001)	0.026*** (0.004)	0.061*** (0.001)	0.060*** (0.001)	0.065*** (0.001)	0.014** (0.004)	0.050*** (0.001)	0.052*** (0.001)	0.054*** (0.001)
White*RM	-0.001 (0.005)	-0.002 (0.002)	0.000 (0.002)	-0.003* (0.001)	0.012* (0.005)	0.010*** (0.002)	0.011*** (0.002)	0.010*** (0.002)	-0.020*** (0.005)	-0.024*** (0.002)	-0.019*** (0.002)	-0.025*** (0.002)
N	2,511,150	16,919,143	16,054,892	16,897,030	2,511,150	16,919,143	16,054,892	16,897,030	2,511,150	16,919,143	16,054,892	16,897,030
Student FE	X	X			X	X			X	X		
Course FE		X	X			X	X			X	X	
Year FE	X	X		X	X	X		X	X	X		X
Student-Year FE			X				X				X	
Student-Course FE				X				X				X

Observations for the *Main* column are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Observations for the *Student FE*, *Student-Course FE*, and *Student-Year FE* are at the student-year-course level for the same years and grades as the main sample. Standard errors clustered at listed level or student-level are shown in parentheses. \* p<0.05, \*\* p<0.01, \*\*\* p<0.001. All columns include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. These coefficients were estimated using equation 1.2 with two changes. First, since these data are analyzed at the course-level, the race match effect captures a direct race match with the teacher. Second, the level of fixed effects were adjusted to match the given label.

**Figure E.1 Race Match Effects by Placebo Race Match Status (Math)**

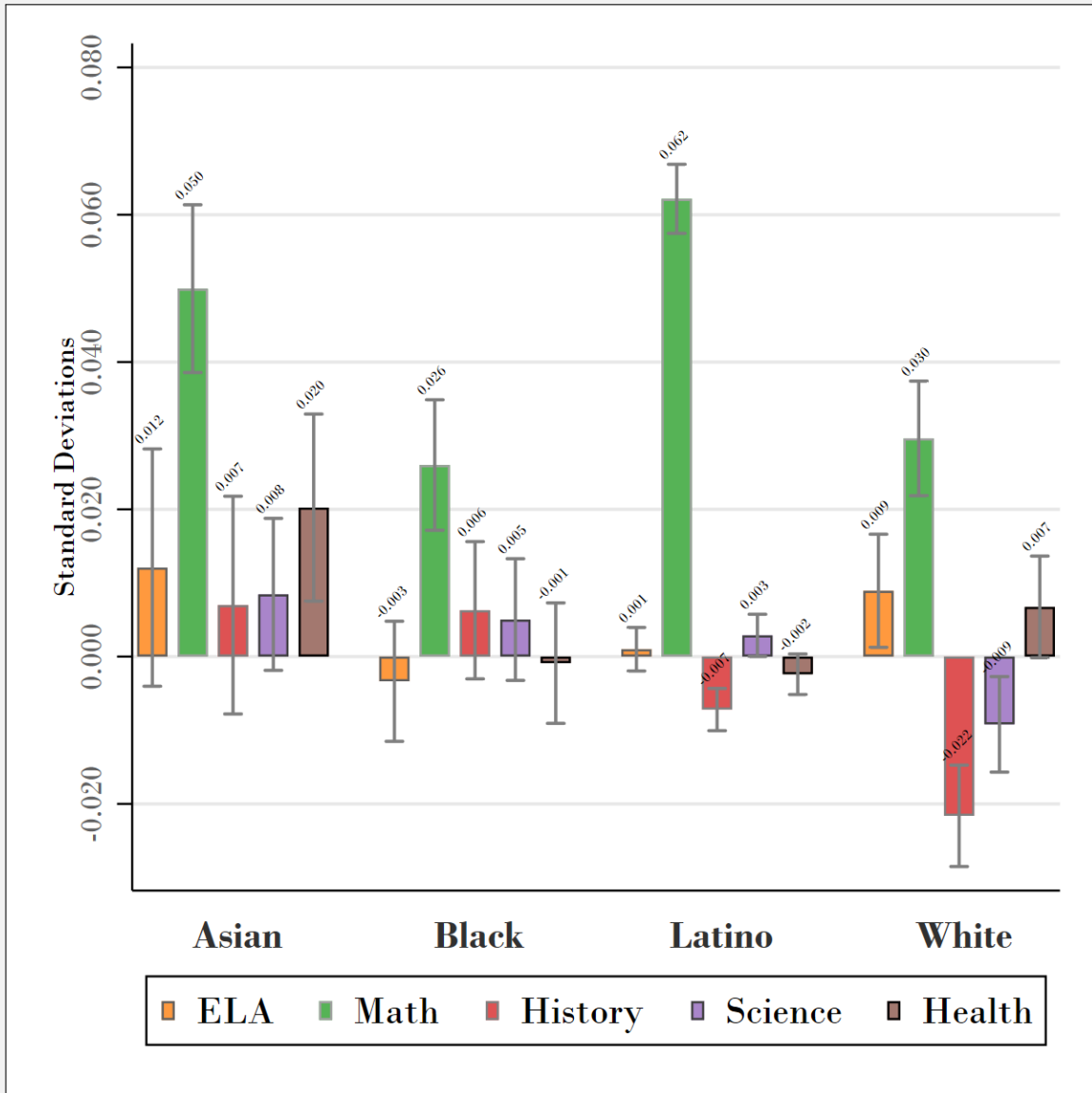
**Panel A: Course Level Placebo Analysis (GPA)**



Placebo sample consists of student-year observations enrolled in each English, Math, History, Science, and Health/P.E. courses within the same year. Observations are restricted from the placebo sample to the subject listed in the title. Each bar represents the coefficient from 1.2 using the race match indicator for the given subject. The titled subject is the estimated race match effect, and the remaining estimates are placebo tests. The whiskers show a 95% confidence interval around each estimate. All regressions include school-grade, student, subject, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the grade-subject-year level. Panel A shows the results for GPA; Panel B for Work Habits; and Panel C for Cooperation.

**Figure E.1 (cont'd) Race Match Effects by Placebo Race Match Status (Math)**

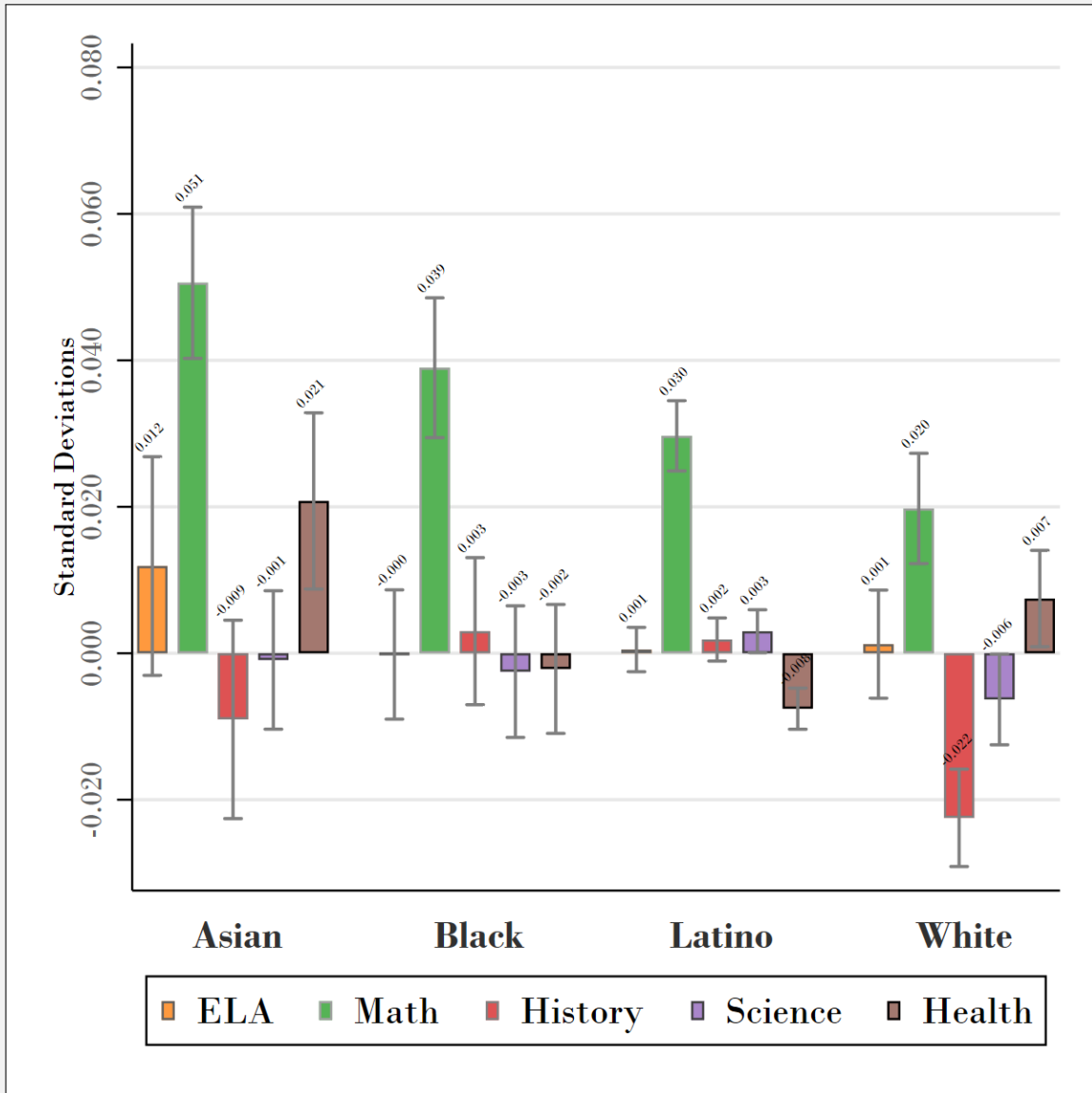
**Panel B: Course Level Placebo Analysis (Work Habits)**



Placebo sample consists of student-year observations enrolled in each English, Math, History, Science, and Health/P.E. courses within the same year. Observations are restricted from the placebo sample to the subject listed in the title. Each bar represents the coefficient from 1.2 using the race match indicator for the given subject. The titled subject is the estimated race match effect, and the remaining estimates are placebo tests. The whiskers show a 95% confidence interval around each estimate. All regressions include school-grade, student, subject, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the grade-subject-year level. Panel A shows the results for GPA; Panel B for Work Habits; and Panel C for Cooperation.

**Figure E.1 (cont'd) Race Match Effects by Placebo Race Match Status (Math)**

**Panel C: Course Level Placebo Analysis (Cooperation)**



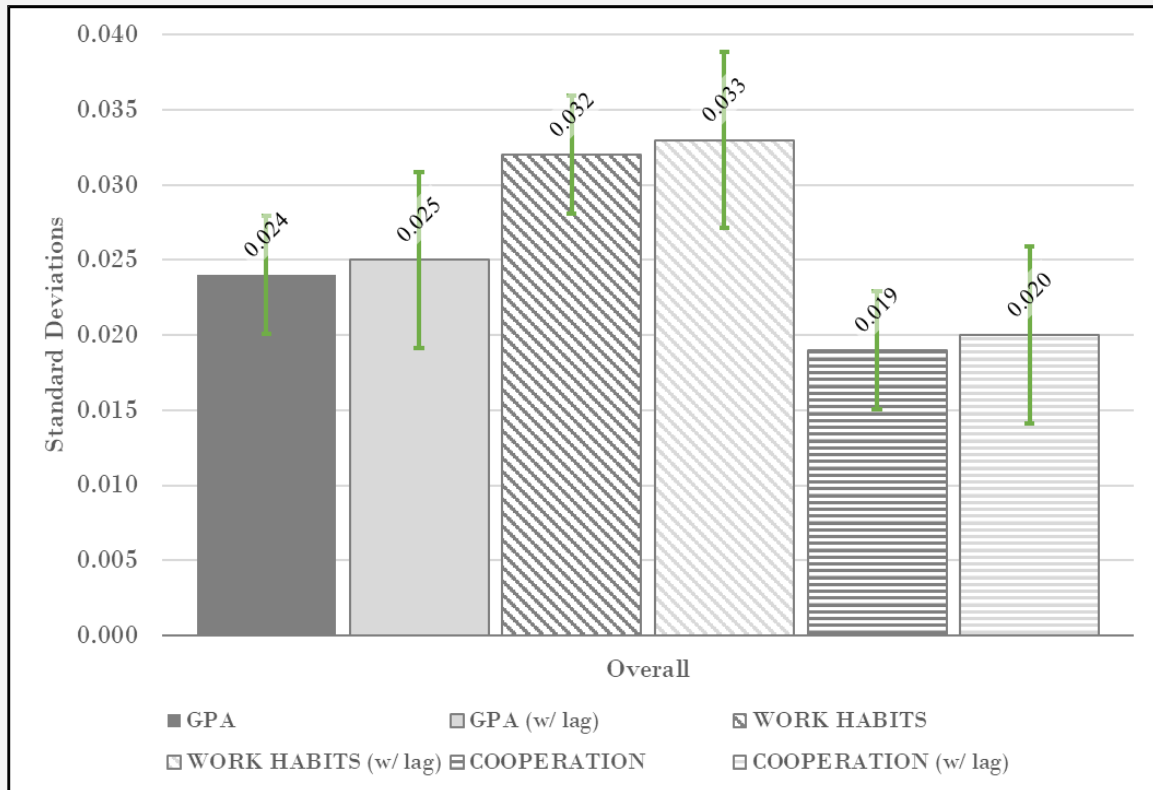
Placebo sample consists of student-year observations enrolled in each English, Math, History, Science, and Health/P.E. courses within the same year. Observations are restricted from the placebo sample to the subject listed in the title. Each bar represents the coefficient from 1.2 using the race match indicator for the given subject. The titled subject is the estimated race match effect, and the remaining estimates are placebo tests. The whiskers show a 95% confidence interval around each estimate. All regressions include school-grade, student, subject, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the grade-subject-year level. Panel A shows the results for GPA; Panel B for Work Habits; and Panel C for Cooperation.



## APPENDIX F

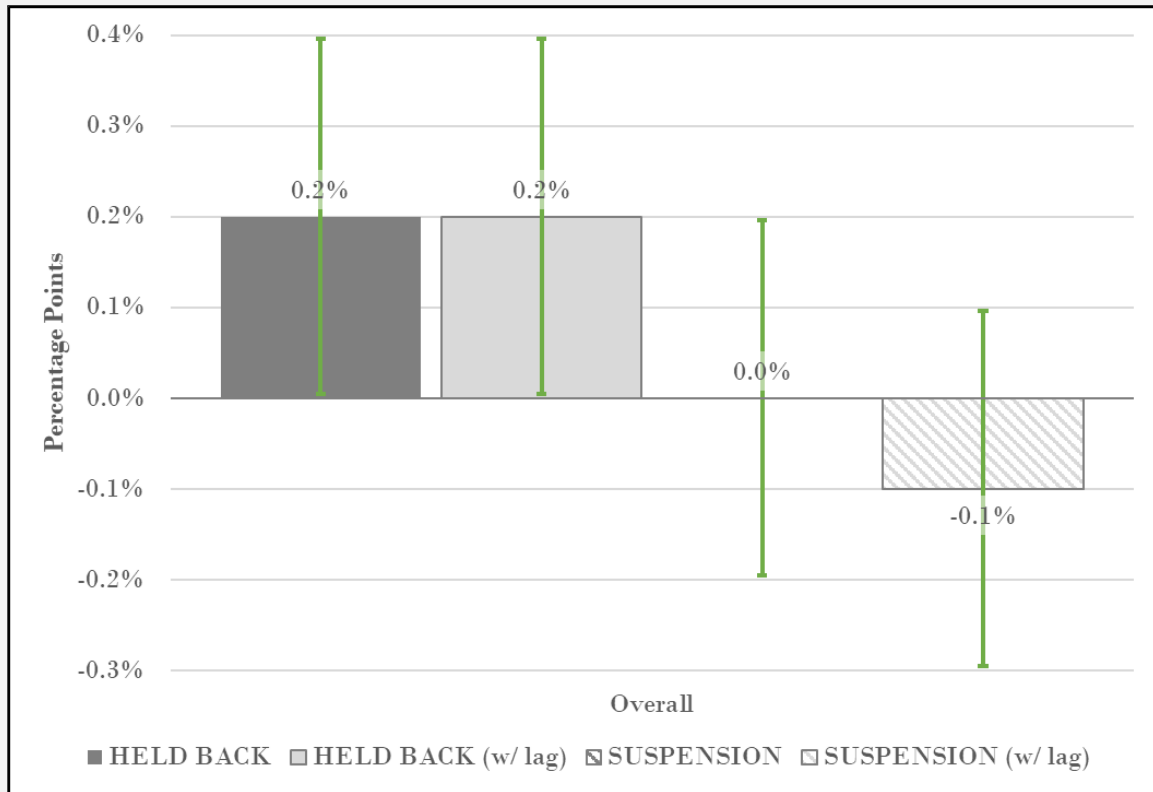
### ROBUSTNESS CHECKS FOR STUDENT-TEACHER RACIAL SORTING WITHIN SCHOOLS

**Figure F.1** Race Match Effect with Lag for Standardized Noncognitive Outcomes



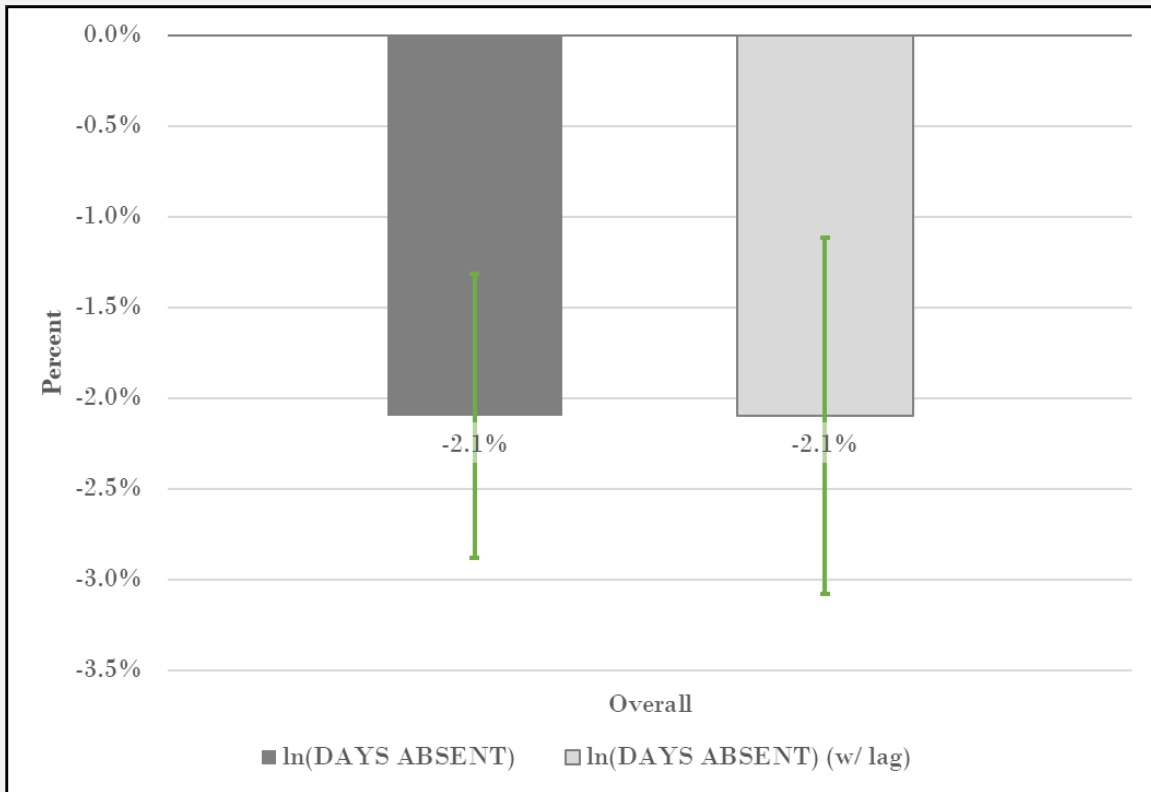
Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered, and the whiskers show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* indicates the results for Equation 1.1. The second coefficient represents the race match effected when including the a lagged outcome into the regression model. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the year level. *Held Back* and *Suspension* are dummy variables and indicate if a student has been held back or suspended for the given year.  $\ln(\text{Days Absent})$  is calculated using  $\ln(x+1)$  to account for zeros. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure is analogous to Table 1.4 in Section 1.6

**Figure F.2 Race Match Effect with Lag for Probability Noncognitive Outcomes**



Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered, and the whiskers show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* indicates the results for Equation 1.1. The second coefficient represents the race match effected when including the a lagged outcome into the regression model. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the year level. *Held Back* and *Suspension* are dummy variables and indicate if a student has been held back or suspended for the given year.  $\ln(\text{Days Absent})$  is calculated using  $\ln(x+1)$  to account for zeros. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure is analogous to Table 1.4 in Section 1.6

**Figure F.3 Race Match Effect with Lag for  $\ln(\text{Days Absent})$**

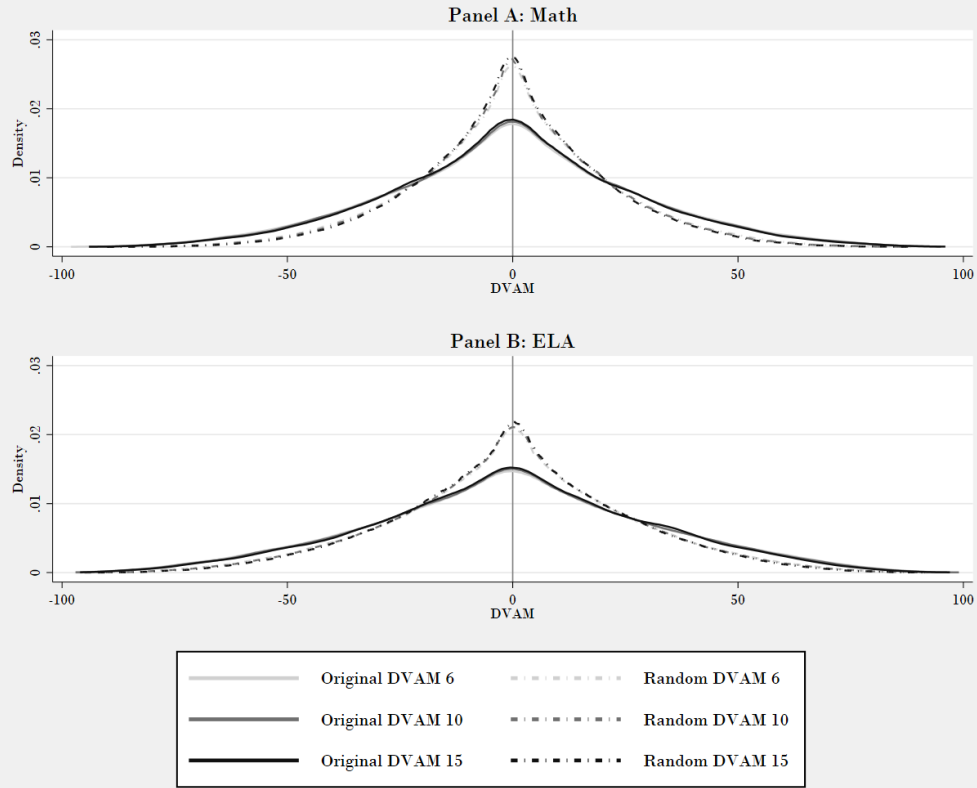


Observations are at the student-year level for school years 2007-08 through 2017-18 for students in grades 6 through 12. Standard errors are clustered at the student-clustered, and the whiskers show the 95% confidence interval. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . *Overall* indicates the results for Equation 1.1. The second coefficient represents the race match effected when including the a lagged outcome into the regression model. Both models include school-grade, student, and year fixed effects and time-variant student, teacher, and class-averaged characteristics. *GPA*, *Work Habits*, and *Cooperation* are standardized to the year level. *Held Back* and *Suspension* are dummy variables and indicate if a student has been held back or suspended for the given year.  $\ln(\text{Days Absent})$  is calculated using  $\ln(x+1)$  to account for zeros. The treatment variable, *Share RM*, is the share of race matched teachers a student has in a given year. This figure is analogous to Table 1.4 in Section 1.6

## APPENDIX G

### ROBUSTNESS CHECK FOR THRESHOLD LEVELS

**Figure G.1** DVAM and Random DVAM Distributions



NOTE: This figure compares the DVAM distribution of our sample to the DVAM distribution generated when randomly assigning student with disability (SWD) status within each classroom, where DVAM represents the difference in a teacher's percentile ranking in SWD VAM from their percentile ranking in non-SWD VAM. A negative DVAM indicates a relative advantage for instructing SWDs. The numbers within the legend signify the threshold samples used in generating the VAMs. For example, *Original DVAM 6* was created using only teachers with at least 6 SWDs and at least 6 non-SWDs across the panel. Since the randomization process preserves the number of SWDs and non-SWDs within each classroom, the *Random DVAMs* uses an identical sample as the *Original DVAM* analogue.

**Table G.1** Association between Student Disability Status and Teacher VAMs by Threshold Sample

Panel A: Math						
	SWD VAM			Non-SWD VAM		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Sample</i>	6	10	15	6	10	15
SWD	-0.036*** (0.005)	-0.036*** (0.005)	-0.035*** (0.006)	-0.042*** (0.004)	-0.043*** (0.004)	-0.041*** (0.005)
N	1,582,059	1,408,173	1,174,409	1,582,059	1,408,173	1,174,409
Panel B: ELA						
	SWD VAM			Non-SWD VAM		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Sample</i>	6	10	15	6	10	15
SWD	-0.044*** (0.006)	-0.040*** (0.006)	-0.037*** (0.006)	-0.056*** (0.004)	-0.056*** (0.004)	-0.054*** (0.005)
N	1,664,316	1,481,418	1,246,446	1,664,316	1,481,418	1,246,446

Threshold samples contain students taught by teachers with at least the given threshold quantity (i.e., 6, 10, or 15) of SWD and non-SWD observations across the panel dataset. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year fixed effects. VAM (value-added measure). SWD (student with disability). ELA (English language arts). \*p<.05. \*\*p<.01. \*\*\*p<.001.

## APPENDIX H

### SWD SORTING BY DISABILITY TYPE

**Table H.1** Association between Student Disability Type and Teacher VAMs

	Math VAM		ELA VAM	
	(1) SWD	(2) Non-SWD	(3) SWD	(4) Non-SWD
Autism	-0.006 (0.007)	0.005 (0.006)	-0.006 (0.008)	-0.012* (0.006)
SLD	-0.041*** (0.006)	-0.053*** (0.005)	-0.051*** (0.007)	-0.065*** (0.005)
LS Imp.	-0.006 (0.005)	-0.001 (0.004)	0.004 (0.005)	0.002 (0.005)
Other Disability	-0.026*** (0.006)	-0.030*** (0.004)	-0.023*** (0.005)	-0.043*** (0.004)
N	1,408,173	1,408,173	1,481,418	1,481,418

Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year fixed effects. SWD (student with disability). VAM (value-added measure). SLD (specific learning disability). LS Imp. (language and speech impairment). \*p<.05. \*\*p<.01. \*\*\*p<.001.

**Table H.2** Association between Student Disability Type and Teacher Relative Advantage

Relative Advantage Definition	Math			ELA		
	(1) >0	(2) >10	(3) >20	(4) >0	(5) >10	(6) >20
Autism	-0.004 (0.004)	-0.003 (0.003)	-0.001 (0.003)	0.003 (0.004)	0.002 (0.004)	-0.000 (0.003)
SLD	0.006* (0.003)	0.006* (0.003)	0.004 (0.002)	0.007* (0.003)	0.009** (0.003)	0.006* (0.003)
LS Imp.	-0.007** (0.002)	-0.003 (0.002)	-0.002 (0.002)	0.002 (0.003)	0.001 (0.002)	-0.001 (0.002)
Other Disability	0.005* (0.002)	0.003 (0.002)	0.003 (0.002)	0.010*** (0.003)	0.008** (0.003)	0.008*** (0.002)
N	1,408,173	1,408,173	1,408,173	1,481,418	1,481,418	1,481,418

Outcome variables are dummy indicators for relative advantage. >0 (>10; >20) indicate that the SWD VAM percentile is more than 0 (10; 20) percentile points greater than the non-SWD VAM percentile. Teacher-clustered standard errors shown in parentheses. All columns include school-grade-year fixed effects. ELA (English language arts). SLD (specific learning disability). LS Imp. (language and speech impairment). VAM (value-added measure). \*p<.05. \*\*p<.01. \*\*\*p<.001.

## APPENDIX I

### RELATIONSHIP BETWEEN VAM/RELATIVE ADVANTAGE AND SWITCHING SCHOOLS WITHIN LAUSD

**Table I.1** VAMs and Relative Advantage Association with Switching Schools within LAUSD

Panel A: VAM						
	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
SWD VAM	-0.0023 (0.0014)		0.0010 (0.0017)	0.0008 (0.0014)		0.0019 (0.0015)
Non-SWD VAM		-0.0057*** (0.0014)	-0.0063*** (0.0016)		-0.0020 (0.0014)	-0.0028 (0.0015)
Panel B: Relative Advantage						
	Math			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Advantage (RA)	0.0047 (0.0028)	0.0069* (0.0030)	0.0069* (0.0034)	0.0035 (0.0028)	0.0056 (0.0030)	-0.0014 (0.0033)
<i>RA Definition</i>	>0	>10	>20	>0	>10	>20
N	33,748	33,748	33,748	34,489	34,489	34,489
Switch school is a dummy indicator that equals 1 when a teacher changes schools within Los Angeles Unified School District the following year. Relative advantage variables are dummy indicators that equal 1 when a teacher has a VAM SWD percentile > VAM non-SWD percentile for a given year in that subject. >10 (>20) indicate that the VAM SWD percentile is more than 10 (20) percentile points greater than VAM non-SWD percentile. Teachers that leave due to reduction in force are excluded. Teacher-clustered standard errors shown in parentheses. All columns include school and year fixed effects. Coefficients for teacher and class characteristics not shown. VAM (value-added measure). SWD (student with disability). ELA (English language arts) RA (relative advantage). *p<.05. **p<.01. ***p<.001.						



## **BIBLIOGRAPHY**

## BIBLIOGRAPHY

- Achinstein, B., Ogawa, R. T., Sexton, D., and Freitas, C. (2010). Retaining teachers of color: A pressing problem and a potential strategy for “hard-to-staff” schools. *Review of educational research*, 80(1):71–107.
- Antrop-González, R. and De Jesus, A. (2006). Toward a Theory of Critical Care in Urban Small School Reform: Examining Structures and Pedagogies of Caring in Two Latino Community-Based Schools. *International Journal of Qualitative Studies in Education*, 19:409–433.
- Barbaranelli, C., Caprara, G., Rabasca, A., and Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences*, 34:645–664.
- Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology. General Section*, 28(1):97–104.
- Blanton, L., Pugach, M., and Boveda, M. (2018). Interrogating the Intersections Between General and Special Education in the History of Teacher Education Reform. *Journal of Teacher Education*, 69(4):354–366.
- Blanton, L., Pugach, M., and Florian, L. (2011). Preparing General Education Teachers to Improve Outcomes for Students with Disabilities. American Association of Colleges for Teacher Education & National Center, Washington, D. C.
- Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2003). The Draw of Home: How Teachers’ Preferences for Proximity Disadvantage Urban Schools. *Journal of Policy Analysis and Management*, 24(1):113–132.
- Brookings Institute (2013). The Resurgence of Ability Grouping and Persistence of Tracking. <https://www.brookings.edu/research/the-resurgence-of-ability-grouping-and-persistence-of-tracking/>.
- Brownell, M., Sindelar, P., Kiely, M., and Danielson, L. (2010). Special Education Teacher Quality and Preparation: Exposing Foundations, Constructing a New Model. *Exceptional Children*, 76(3):357–377.
- Buddin, R. and Zamarro, G. (2009a). Teacher qualifications and middle school student achievement.
- Buddin, R. and Zamarro, G. (2009b). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics*, 66(2):103–115.
- Cherng, H.-Y. S. and Halpin, P. F. (2016). The Importance of Minority Teachers: Student Percep-

- tions of Minority Versus White Teachers. *Educational Researcher*, 45(7):407–420.
- Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9):2593–2632.
- Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014b). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–79.
- Child, A., Cirino, P., Fletcher, J., Willcutt, E., and Fuchs, L. (2019). A Cognitive Dimensional Approach To Understanding Shared And Unique Contributions To Reading, Math and Attention Skills. *Journal Of Learning Disabilities*, 52(1):15–30.
- Chin, M. J., Quinn, D. M., Dhaliwal, T. K., and Lovison, V. S. (2020). Bias in the Air: A Nationwide Exploration of Teachers’ Implicit Racial Attitudes, Aggregate Bias and Student Outcomes. *Educational Researcher*, 49(8):566–578.
- Chingos, M. and Peterson, P. (2011). It’s Easier To Pick A Good Teacher Than To Train One: Familiar And New Results On The Correlates Of Teacher Effectiveness. *Economics Of Education Review*, 30(3):449–465.
- Chudowsky, N. and Chudowsky, V. (2009). State Test Score Trends through 2007-08, Part 4: Has Progress Been Made in Raising Achievement for Students with Disabilities? *Center on Education Policy*.
- Connor, C., Morrison, F., Fishman, B., Giuliani, S., Luck, M., Underwood, P., Bayraktar, A., Crowe, E., and Schatschneider, C. (2011). Testing the impact of child characteristics  $\times$  instruction interactions on third graders’ reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, 46(3):189–221.
- Connor, C. M., Phillips, B. M., Kim, Y.-S. G., Lonigan, C. J., Kaschak, M. P., Crowe, E., Dombek, J., and Al Otaiba, S. (2018). Examining the Efficacy of Targeted Component Interventions on Language and Literacy for Third and Fourth Graders Who are at Risk of Comprehension Difficulties. *Scientific Studies of Reading*, 22(6):462–484.
- Cook, B. (2002). Inclusive Attitudes, Strengths and Weaknesses of Pre-Service General Educators Enrolled in a Curriculum Infusion Teacher Preparation Program. *Teacher Education And Special Education*, 25(3):262–277.
- Darling-Hammond, L. (1998). Unequal opportunity: Race and education.
- Dee, T. S. (2004). Teachers, Race and Student Achievement in a Randomized Experiment. *The Review of Economics and Statistics*, 86(1):195–210.

- Dee, T. S. (2005). A Teacher like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review*, 95(2):158–165.
- Diamond, J. B., Randolph, A., and Spillane, J. P. (2004). Teachers' Expectations and Sense of Responsibility for Student Learning: The Importance of Race, Class and Organizational Habitus. *Anthropology & Education Quarterly*, 35(1):75–98.
- Downey, D. and Pribesh, S. (2004). When Race Matters: Teachers' Evaluations of Students' Classroom Behavior. *Sociology of Education*, 77.
- Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: Perseverance and Passion for Long-Term Goals. *Journal of Personality and Social Psychology*, 92(6):1087–1101.
- EdSource (2017). Q & A: Poor students and English learners did better on tests but achievement gaps yet to narrow. <https://edsource.org/2017/q-a-state-test-achievement-gap-expert-says-poor-students-english-learners-falling-further-behind/583017>.
- Egalite, A. J., Kisida, B., and Winters, M. A. (2015). Representation in the Classroom: The Effect of Own-race Teachers on Student Achievement. *Economics of Education Review*, 45:44–52.
- Ehrenberg, R. G., Goldhaber, D. D., and Brewer, D. J. (1995). Do Teachers' Race, Gender and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review*, 48(3):547–561.
- Elder, T. and Zhou, Y. (2021). The Black-White Gap in Noncognitive Skills among Elementary School Children. *American Economic Journal: Applied Economics*, 13(1):105–32.
- Elstad, E., Christophersen, K.-A., and Turmo, A. (2020). Preservice Teachers' Peer Support Through Close Social Ties: An Investigation Of The Social Side Of Teacher Education. *Advances in Education Sciences*, 2(2):4–28.
- Endrew F. vs. Douglas County School District RE-1 (2017). 580 U.S \_\_\_\_.
- Ferguson, R. F. (2003). Teachers' Perceptions and Expectations and the Black-White Test Score Gap. *Urban Education*, 38(4):460–507.
- Florian, L. (2012). Preparing Teachers To Work In Inclusive Classrooms: Key Lessons For The Professional Development Of Teacher Educators From Scotland's Inclusive Practice Project. *Journal Of Teacher Education*, 63(4):275–285.
- Fox, L. (2015). Seeing Potential: The Effects of Student–Teacher Demographic Congruence on Teacher Expectations and Recommendations. *AERA open*, 2(1):2332858415623758.

- Frank, K. A., Lo, Y.-J., and Sun, M. (2014). Social network analysis of the influences of educational reforms on teachers' practices and interactions. *Zeitschrift für Erziehungswissenschaft*, 17(5):117–134.
- Fryer, Roland G., J. and Levitt, S. D. (2013). Testing for Racial Differences in the Mental Ability of Young Children. *American Economic Review*, 103(2):981–1005.
- Fuchs, L., Geary, D., Fuchs, D., Compton, D., and Hamlett, C. (2016). Pathways To Third-Grade Calculation Versus Word-Reading Competence: Are They More Alike or Different? *Child Development*, 87(2):558–567.
- Galiatsos, S., Kruse, L., and Whittaker, M. (2019). Forward Together: Helping Educators Unlock the Power of Students Who Learn Differently. *National Center for Learning Disabilities*.
- Gay, G. (2000). *Culturally Responsive Teaching: Theory, Research and Practice*. Teachers College Press, 1st edition.
- Gershenson, S., Hart, C. M. D., Hymna, J., Lindsay, C. A., and Papageorge, N. W. (2018). The Long-Run Impact of Same-Race Teachers. *National Bureau of Economic Research*, (Working Paper 25254).
- Gershenson, S., Holt, S. B., and Papageorge, N. W. (2016). Who Believes in Me? The Effect of Student-Teacher Demographic Match on Teacher Expectations. *Economics of Education Review*, 52:209–224.
- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J., and Witzel, B. (2009a). Assisting students struggling with mathematics: Response to intervention (rti) for elementary and middle schools. pages (NCEE 2009–4060). Washington, DC: National Center for Education Evaluation and Regional Services, Institute of Education Sciences, U.S. Department of Education.
- Gersten, R., Chard, D., Jayanthi, M., Baker, S., Morphy, P., and Flojo, J. (2009b). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79:1202–1242.
- Goings, R. B. and Bianco, M. (2016). It's Hard to Be Who You Don't See: An Exploration of Black Male High School Students' Perspectives on Becoming Teachers. *Urban Review: Issues and Ideas in Public Education*, 48(4):628–646.
- Goldhaber, D., Gross, B., and Player, D. (2011). Teacher Career Paths, Teacher Quality and Persistence In The Classroom: Are Public Schools Keeping Their Best? *J. Pol. Anal. Manage*, 30:57–87.
- Grissom, J. A., Kern, E. C., and Rodriguez, L. A. (2015). The "Representative Bureaucracy" in Education: Educator Workforce Diversity, Policy Outputs and Outcomes for Disadvantaged

- Students. *Educational Researcher*, 44(3):185–192.
- Grissom, J. A., Nicholson-Crotty, J., and Nicholson-Crotty, S. (2009). Race, region and representative bureaucracy. *Public Administration Review*, 69(5):911–919.
- Grissom, J. A., Rodriguez, L., and Kern, E. (2017). Teacher and Principal Diversity and the Representation of Students of Color in Gifted Programs. *The Elementary School Journal*, 117(3):396–422.
- Guskey, Thomas R.; Link, L. J. (2018). Exploring the Factors Teachers Consider in Determining Students’ Grades. *Assessment in Education: Principles, Policy & Practice*, 26(3):303–320.
- Hanushek, E., Rivkin, S., and Schiman, J. (2016). Dynamic Effects of Teacher Turnover on the Quality of Instruction. *Economics of Education Review*, 54:132–148.
- Hanushek, E. A. (2004). Why Public Schools Lose Teachers. *Journal of Human Resources*, 39(2):326–354.
- Harbatkin, E. (2021). Does student-teacher race match affect course grades? *Economics of Education Review*, 81.
- Hashim, A., Strunk, K., and Dhaliwal, T. (2018). Justice for All? Suspension Bans and Restorative Justice Programs in the Los Angeles Unified School District. *Peabody Journal of Education*, 93.
- Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103(6):2052–86.
- Heckman, J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.
- Howley, A., Kusimo, P. S., and Parrott, L. (2000). Grading and the Ethos of Effort. *Learning Environments Research*, 3(3):229–246.
- Huddleston, A. P. (2014). Achievement at whose expense? a literature review of test-based grade retention policies in us schools. *EducationPolicy Analysis Archives*, 22(18).
- Institute, A. S. (2015). *The state of teacher diversity in American education*. ERIC Clearinghouse.
- Irvine, J. J. (1989). Beyond Role Models: An Examination of Cultural Influences on the Pedagogical Perspectives of Black Teachers. *Peabody Journal of Education*, 66(4):51–63.

- Jackson, C. K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy*, 126(5):2072–2107.
- Jackson, C. K. and Bruegmann, E. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, 1(4):85–108.
- Jennings, J. L. and DiPrete, T. A. (2010). Teacher effects on social and behavioral skills in early elementary school. *Sociology of Education*, 83(2):135–159.
- Jones, N. and Winters, M. A. (2022). Are two teachers better than one? the effect of co-teaching on students with and without disabilities. *Journal of Human Resources*, 0420-10834R3.
- Jones, N. D., Bell, C. A., Brownell, M., Qi, Y., Peyton, D., Pua, D., Fowler, M., and Holtzman, S. (2022). Using classroom observations in the evaluation of special education teachers. *Educational Evaluation and Policy Analysis*.
- Joshi, E., Doan, S., and Springer, M. G. (2018). Student-Teacher Race Congruence: New Evidence and Insight From Tennessee. *AERA Open*, 4(4).
- Kamens, M., Loprete, S., and Slostad, F. (2000). Classroom Teachers' Perceptions about Inclusion and Preservice Teacher Education. *Teaching Education*, 11(2):147–158.
- Kane, T. J., Mccaffrey, D. F., Miller, T., and Staiger, D. O. (2013). Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. In *Research Paper. MET Project. Bill & Melinda Gates Foundation*.
- Kane, T. J. and Staiger, D. O. (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. *National Bureau of Economic Research*, December(Working Paper 14607).
- Kettler, T. and Hurst, L. T. (2017). Advanced Academic Participation: A Longitudinal Analysis of Ethnicity Gaps in Suburban Schools. *Journal for the Education of the Gifted*, 40(1):3–19.
- King, S. H. (1993). The limited presence of african-american teachers. *Review of Educational Research*, 63(2):115–149.
- Koedel, C., Mihaly, K., and Rockoff, J. (2015). Value-Added Modeling: A Review. *Economics Of Education Review*, 47:180–195.
- Ladson-Billings, G. (1995). Toward a Theory of Culturally Relevant Pedagogy. *American Educational Research Journal*, 32(3):465–491.
- Ladson-Billings, G. (2004). *Crossing over to Canaan: The journey of new teachers in diverse classrooms*. John Wiley & Sons.

- LAUSD (2018). *A Parent's Guide to Special Education Services*. Los Angeles Unified School District. <https://tinyurl.com/3x94x743>.
- LAUSD (2022). *Fingertip Facts 2021-2022*. Los Angeles Unified School District.
- Lim, H.-H. (2006). Representative bureaucracy: Rethinking substantive effects and active representation. *Public Administration Review*, 66(2):193–204.
- Lindsay, C. and Hart, C. (2017). Exposure to Same-Race Teachers and Student Disciplinary Outcomes for Black Students in North Carolina. *Educational Evaluation and Policy Analysis*, 39(3):1–9.
- Liu, J., Hayes, M. S., and Gershenson, S. (2022). From Referrals to Suspensions: New Evidence on Racial Disparities in Exclusionary Discipline. *Annenberg Working Paper*, January (EdWorkingPaper No. 21-442).
- Liu, J. and Loeb, S. (2021). Engaging Teachers: Measuring the Impact of Teachers on Student Attendance in Secondary School. *Journal of Human Resources*, 56(2):N.PAG.
- Lockwood, J. and McCaffrey, D. (2014). Correcting For Test Score Measurement Error In Ancova Models For Estimating Treatment Effects. *Journal Of Educational And Behavioral Statistics*, 39(1):22–52.
- Loeb, S., Soland, J., and Fox, L. (2014). Is A Good Teacher A Good Teacher For All? Comparing Value-Added Of Teachers With Their English Learners And Non-English Learners. *Educational Evaluation And Policy Analysis*, 36(4):457–475.
- Lopez, A. E. (2016). *Culturally Responsive and Socially Just Leadership in Diverse Contexts: From Theory to Action*. Palgrave Macmillan.
- Los Angeles Almanac (2021). Asian Ethnic Origin Los Angeles County: 2010 Census & 2019 Census Estimates. <http://www.laalmanac.com/population/po16.php>.
- Master, B., Loeb, S., Whitney, C., and Wyckoff, J. (2016). Different Skills?: Identifying Differentially Effective Teachers of English Language Learners. *The Elementary School Journal*, 117(2):261–284.
- Mawhinney, L. (2010). Let's lunch and learn: Professional knowledge sharing in teachers' lounges and other congregational spaces. *Teaching and Teacher Education*, 26(4):972 – 978.
- McGrady, P. B. and Reynolds, J. R. (2013). Racial Mismatch in the Classroom: Beyond Black-white Differences. *Sociology of Education*, 86(1):3–17.
- Meier, K. J. and Stewart Jr, J. (1992). The impact of representative bureaucracies: Educational systems and public policies. *The American Review of Public Administration*, 22(3):157–171.



- Milner, H. (2011). Culturally Relevant Pedagogy in a Diverse Urban Classroom. *The Urban Review*, 43:66–89.
- Milner, H. R. (2006). Preservice teachers' learning about cultural and racial diversity: Implications for urban education. *Urban Education*, 41(4):343–375.
- Morris-Mathews, H., Stark, K. R., Jones, N. D., Brownell, M. T., and Bell, C. A. (2021). Danielson's framework for teaching: Convergence and divergence with conceptions of effectiveness in special education. *Journal of Learning Disabilities*, 54(1):66–78.
- NCES (2017). *Digest of Education Statistics: 2017* (NCES 2017- 070). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics.
- NCES (2020). *Statistical analysis report: Higher education* (NCES 97-584). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics.
- Newcomer, S. N. (2018). Investigating the Power of Authentically Caring Student-Teacher Relationships for Latinx Students. *Journal of Latinos and Education*, 17(2):179–193.
- Nicholson-Crotty, J., Grissom, J. A., and Nicholson-Crotty, S. (2011). Bureaucratic Representation, Distributional Equity and Democratic Values in the Administration of Public Programs. *The Journal of Politics*, 73(2):582–596.
- Oates, G. L. S. C. (2003). Teacher-Student Racial Congruence, Teacher Perceptions and Test Performance. *Social Science Quarterly*, 84(3):508–525.
- Office of Planning, Evaluation and Policy Development (2016). U.S. Department of Education, Office of Planning, Evaluation and Policy Development, The State of Racial Diversity in the Educator Workforce.
- Okonofua, J. A., Walton, G. M., and Eberhardt, J. L. (2016). A Vicious Cycle: A Social-Psychological Account of Extreme Racial Disparities in School Discipline. *Perspectives on Psychological Science*, 11(3):381–398.
- Petek, N. and Pope, N. (2021). The Multidimensional Impact of Teachers on Students. *Working Paper*.
- Pew Research Center (2014). Demographic and Economic Profiles of Hispanics by State and County, 2014: California. <https://www.pewresearch.org/hispanic/states/state/ca/>.
- Pew Research Center (2018). What Unites and Divides Urban, Suburban and Rural Communities. <https://www.pewresearch.org/social-trends/2018/05/22/what-unites-and-divides-urban-suburban-and-rural-communities/>.
- Pew Research Center (2021). America's public school teachers are far less racially and ethni-

cally diverse than their students. <https://www.pewresearch.org/fact-tank/2021/12/10/americas-public-school-teachers-are-far-less-racially-and-ethnically-diverse-than-their-students/>.

- Price, J. (2010). The Effect of Instructor Race and Gender on Student Persistence in STEM Fields. *Economics of Education Review*, 29(6):901 – 910.
- Ready, D. D. and Wright, D. L. (2011). 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.
- Reardon, S. F. and Galindo, C. (2009). The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades. *American Educational Research Journal*, 46(3):853–891.
- Roch, C. H., Pitts, D. W., and Navarro, I. (2010). Representative bureaucracy and policy tools: Ethnicity, student discipline and representation in public schools. *Administration & Society*, 42(1):38–65.
- Rocha, R. R. and Hawes, D. P. (2009). Racial diversity, representative bureaucracy and equity in multiracial school districts. *Social Science Quarterly*, 90(2):326–344.
- Rodriguez, G. (2020). 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.
- Sadler, J. (2005). Knowledge, Attitudes and Beliefs of the Mainstream Teachers of Children with a Preschool Diagnosis of Speech/Language Impairment. *Child Language Teaching And Therapy*, 21(2):147–163.
- Saft, E. and Pianta, R. (2001). Teachers' Perceptions of Their Relationships with Students: Effects of Child Age, Gender and Ethnicity of Teachers and Children. *School Psychology Quarterly*, 16(2):125–141.
- Schulte, A. C., Stevens, J. J., Elliott, S. N., and Tindal, G. and Nese, J. F. T. (2016). Achievement Gaps for Students with Disabilities: Stable, Widening, or Narrowing on a State-Wide Reading Comprehension Test? *Journal of Educational Psychology*, 108(7):925–942.
- Shi, Y. and Zhu, M. (2021). Equal Time for Equal Crime? Racial Bias in School Discipline. *IZA Discussion Paper No. 14306*, pages 613–629.
- Shirrell, M., Bristol, T., and Britton, T. (2021). 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).
- Sindelar, P., Brownell, M., and Billingsley, B. (2010). Special Education Teacher Education

- Research: Current Status and Future Directions. *Teacher Education And Special Education*, 33(1):8–24.
- Skiba, R., Horner, R., Chung, C.-G., Rausch, M., May, S., and Tobin, T. (2011). Race Is Not Neutral: A National Investigation of African American and Latino Disproportionality in School Discipline. *School Psychology Review*, 40.
- Spillane, J. P. (2012). *Distributed leadership*, volume 4. John Wiley & Sons.
- Steele, C. M. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *American Psychologist*, 69(5):613–629.
- Steele, C. M. and Aronson, J. (1997). A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. *Journal of Personality and Social Psychology*, 52(6):797–811.
- Stiefel, L., Gottfried, M., Shiferaw, M., and Schwartz, A. (2021). Is Special Education Improving? Case Evidence From New York City. *Journal Of Disability Policy Studies*, 32(2):95–107.
- Stockard, J., Wood, T. W., Coughlin, C., and Rasplika Khoury, C. (2018). The Effectiveness of Direct Instruction Curricula: A Meta-Analysis of a Half Century of Research. *Review of Educational Research*, 88(4):479–507.
- Suhr, D. D. (2005). Principal component analysis vs. exploratory factor analysis. *SAS Conference Proceedings: SAS Users Group International*, 30(203).
- Sun, M., Loeb, S., and Grissom, J. (2017). Building Teacher Teams: Evidence of Positive Spillovers From More Effective Colleagues. *Educational Evaluation and Policy Analysis*, 39(1):104–125.
- Sun, M., Penuel, W. R., Frank, K. A., Gallagher, H. A., and Youngs, P. (2013). Shaping professional development to promote the diffusion of instructional expertise among teachers. *Educational Evaluation and Policy Analysis*, 35(3):344–369.
- Swaak, T. (2020). For the first time in more than 20 years, LAUSD is in full control of its special ed system. As parents worry about accountability, the district shifts its focus. <https://laschoolreport.com/for-the-first-time-in-20-years-laUSD-is-in-full-control-of-its-special-ed-system-as-parents-worry-about-accountability-the-district-shifts-its-focus/>.
- Tenenbaum, H. R. and Ruck, M. D. (2007). Are Teachers' Expectations Different for Racial Minority Than for European American Students? A Meta-Analysis. *Journal of Educational Psychology*, 99(2):253–273.
- Todd, P. and Wolpin, K. (2007). The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps. *Journal of Human Capital*, 1(1):91–136.

- Torgesen, J., Alexander, A., Wagner, R., Rashotte, C., Voeller, K., and Conway, T. (2001). Intensive Remedial Instruction For Children With Severe Reading Disabilities: Immediate And Long-Term Outcomes From Two Instructional Approaches. *Journal Of Learning Disabilities*, 34(1):33–58.
- U.S. Department of Education, National Center for Education Statistics (2018). Table 204.60: Percentage distribution of students 6 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by educational environment and type of disability: Selected years, fall 1989 through fall 2017. In U.S. Department of Education, National Center for Education Statistics (Ed.), *Digest of Education Statistics* (2018 ed.). Retrieved from <https://tinyurl.com/f4tpvsmu>.
- Welsh, R. O. and Little, S. (2018). The school discipline dilemma: A comprehensive review of disparities and alternative approaches. *Review of Educational Research*, 88(5):752–794.
- Wood, W. J. (2022). The Student-Teacher Race Match Effect on Learning Skills and Behavioral Outcomes. [Doctoral dissertation, Michigan State University].
- Wood, W. J. and Lai, I. (2022). The Effect of Same Race Teachers and Faculty on Student Test Scores. [Doctoral dissertation, Michigan State University].
- Wu, Z., Spreckelsen, T. F., and Cohen, G. L. (2021). A meta-analysis of the effect of values affirmation on academic achievement. *Journal of Social Issues*, 77(3):702–750.