Heterogeneous Impact of Teacher-Student Demographic Mismatch on Math and Science Achievement

Md Ohiul Islam[1]

Abstract

This paper focuses on the impact of teacher-student demographic mismatch on student success at 9th grade in U.S. high schools. When students, particularly those of disadvantaged backgrounds, are assigned to teachers with different racial and/or gender identities, they may become subject to the "Golem effect"; lower expectations and biases the teachers may have. In this paper, using restricted-access data from the High School Longitudinal Survey of 2009 (HSLS:09), I investigate whether demographic mismatch between teachers and students in high schools has a negative impact on achievement. HSLS:09 is the only nationally representative high school-level survey data that matches math and science teachers to students and their academic performance. The data allows me to analyze within-student variations of achievement and employ student fixed-effects to identify the effect of demographic mismatch with their teachers. Teacher-student matching and standardized course GPA vary between math and science courses. I do not find any evidence of differential sorting or merit-based math/science course selection or race-based teacher-student matching. I find consistent evidence that having a different-sex teacher is disadvantageous for students of all racial backgrounds. Having a different-sex and different-race teacher is associated with achievement loss (0.3 standard deviations), especially for Black female students. The reference matching category is a 9th-grader being assigned to a same-sex and same-race teacher. The importance of studying the effect of demographic mismatch on math-/science outcomes lies in the fact that high-school math and science courses are foundations of future tertiary STEM education and that demographic test score gaps are transmitted to post-secondary levels.

JEL Codes: I21, I29, J15

[1]Postdoctoral Instructor at Economics Department, University of Nevada-Reno. Email: oislam@unr.edu; mdohiul.islam@wmich.edu

1 Introduction

This paper examines the impact of the demographic mismatch between math and science teachers and 9th-grade students in U.S. high schools. Math and science education at high school level is foundational for college-level specialization in STEM fields. STEM stands for Science, Technology, Engineering, and Mathematics. STEM is meant to express a multidisciplinary curriculum focused on scientific fields according to "Digest of Education Statistics. Pocket Digest" (n.d.). Studying early student success in STEM courses at the high school level is important because majoring and minoring in STEM fields at the graduate and undergraduate levels would define career and lifetime earning paths for individuals (Deming and Noray, 2018; Deming and Noray, 2020; Winters, 2014 for review) and the high-school level preparation in math and science serves as the foundation of college-level STEM education. Moreover, labor market outcomes are connected with pre-college and college-level educational experiences and outcomes in different pathways (Altonji, 1992; Loury and Garman, 1995).

Pre-college academic success is driven by a myriad of factors including teachers' effectiveness, peer-to-peer teaching, teaching practices, school curriculum, classroom size, teachers' gender and racial identity, etc. (see Lavy, 2016; Kimbrough et al., 2017; Sansone, 2017; Hanushek, 1999). Teachers' motivational input to students has been credited to be a major channel that can positively improve student achievement in pre-college education (Feldman, 1988). A strand of economic literature studied how students may see same-race and same-gender teachers as role models and feel motivated to improve academic outcomes (see Gershenson et al., 2018; Dee, 2004, Ehrenberg and Brewer, 1995; Sansone, 2017). Economists believe that one of the channels through which race-matching works is creating "role-model" effects. The role-model effect refers to the positive impact a minority teacher may have on a student from a low-income background and/or minority community. Presence of a role-model teacher in class increase a minority student's effort, confidence, and enthusiasm (King, 1993; Cziek, 1995). The opposite of the role-model effect or Pygmalion effect is known as the Golem effect which can negatively affect student achievement. Why pedagogy under racial matching improves student outcomes is a contentious subject but economists credit common life experiences and common cultural background for it. Some also mention that same-race and same-gender teachers may be seen as role models by racial minority students and female students (Dee, 2005; Gershenson et al., 2018; Hoffmann and Oreopoulos, 2009). This paper attempts to analyze the negative effect of demographic mismatch or having a different race and different-sex on STEM achievement of high-school students following the footsteps of Dee (2004).

In this paper, I use a nationally representative data from the High School Longitudinal Study of 2009 to test whether demographic mismatch between students and teachers systemically affects student success[2]. In this survey data, two teachers math and science are included allowing the researcher to purge observable and unobserved factors that are connected to both demographic mismatch and student achievement. Moreover, this data captures the variation in the demographic mismatch between math and science courses in 9th-grade. I test for how students are sorted into classes and do not find any proof of differential sorting.

This data captures the variation in student outcomes and variation in demographic identities of different teachers across courses. This quality allows the researcher to look into the within-student effect of demographic mismatch. In the ideal case, demographic match or mismatch should have no effect on student outcomes. If within-student differences in teacher demographic identities are systemically related to student outcomes, then it would mean demographic mismatch is disadvantageous for a subset of students that experiences the negative impact of it. For the main empirical analysis in this paper, I reshape the data to give it a panel structure, where the student ID is the panel variable. This allows student outcome to vary between math and science GPA in 9th grade. It also allows demographic mismatch to vary between two courses. Variation of GPA between science and math due to differences in demographic match/mismatch can be analyzed in a student-specific panel model, which also allows controlling for the student-specific observed and unobserved effects with student fixed effects. This identification strategy is employed

[2]Source: *High School Longitudinal Study of 2009 (HSLS:09/16), Base-year Survey* (2009)

by Dee (2004), Gershenson et al. (2016), Fairlie et al. (2014) and many others. The primary student achievement variable in this paper is the math and science GPA in 9th-grade z-scored at the school level.

Dee (2004) and Gershenson et al. (2016) look at the teachers' expectation as an outcome variable in their research. Expectations are inherently noisy. If the noise is not random and is systemically connected to the main explanatory variables, the estimated results would be biased (Mullainathan and Bertrand, 2001). Especially, Gershenson et al. (2016) builds a model with student fixed effects and not with teacher fixed effects[3]. If a teacher is consistent with their bias, then the teacher fixed-effect model may at least mitigate the problem of the expectational outcome being noisy. Unlike these authors, I only look at student outcomes that are measured and published by the schools. 9th grade GPA data comes from student certificates that HSLS:09 has collected and arranged. I compile this grade data of the sampled students and calculate their GPA using the formula HSLS:09 has published[4]. Moreover, I test for possible selection bias arising from merit-based placement in math and science classes which Gershenson et al. (2016) and Sansone (2017) did not conduct. I do not find any evidence that students on average received merit-based placement in 9th-grade classes. Furthermore, in 9th-grade students are offered required and/or foundational courses in math and science limiting their choice of course and self-selection into these courses. When not corrected or adjusted for, selection bias can damage the validity of the estimated parameters.

The primary contribution of this paper broadening our perspective on the impact of demographic mismatch on student outcomes and accounting for multiple sources of heterogeneity in the mismatch effect across courses and within a school. Moreover, the paper shows robust and consistent evidence of the negative effect on STEM GPA in 9th grade when a student is assigned to a different-sex teacher. The

[3]Gershenson et al. (2016) used data from the Educational Longitudinal Survey of 2002 in their paper. ELS:2002 asks Math and English teachers how far they think the student would go (academically). The answer options include (a) finish high school, (b) enter college but not finish college, (c) finish college, etc. Using information from the follow-up waves the authors were able to determine if demographic mismatch played a negative role in teachers' expectation formation. They find the non-Black teachers were pessimistic about the students' academic achievement.

[4]The 9th grade GPA is calculated using course-specific information from a restricted-access file provided by the IES. Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of2009 (HSLS:09/16), "Base-year Survey, 2009."

evidence of the negative effect of having a different-race teacher mainly exists for Asian students and Black female students. The latter is important for policymakers as a racial gap in educational attainment is well-documented (Ferguson, 2003; Arcidiacono et al., 2012; Fryer Jr. and Katz, 2013). There are three categories of demographic mismatch discussed in this paper: different sex and different race, different race and same sex, different sex and same race. For example, a teacher is Asian and the student is Black, and both are female, then different race and same sex = 1. Same sex and same race is the reference category, compared to which the effect of demographic mismatch are measured. Having a Same-sex and same-race could be the source of role model or pygmalion effect and would generate positive impacts on student outcomes. This paper proceeds as follows: Section 2 briefly discusses some relevant theoretical and empirical works on the channel of demographic match and mismatch. Section 3 discusses the data. Section 4 describes the main model and identification method. Section 5 discusses necessary tests for the validity of the model parameters of interest. Section 6 discusses empirical results and Section 7 concludes.

2 Literature Review

Teachers affect student outcomes in many ways. Teachers' expectations, encouragement, persistence, and motivation can be broadly defined as subjective inputs which can be correlated to the demographic identities of the teacher and his/her students. Teachers' heightened expectations can be self-fulfilling for students. This effect, known as the "Pygmalion effect" has been studied thoroughly and experimented with. Rosenthal (1968) conducted a field experiment in a US public elementary school that showed heightened teacher expectation was associated with faster IQ gains. Eden and Shani (1982) conducted the same experiment on members of Israeli Defense Forces. They concluded that leadership behavior was a key mediator in generating the Pygmalion effect. Jussim and Harber (2005) find that the overall effect existed, but it was nominal. Bertrand (2016) includes a survey of the literature on the Pygmalion effect and Golem effect. The self-fulfilling nature of lowering expectations is known as the "Golem effect" (Reynolds, 2007). An example of the

Golem effect has been found in Rist (1970) who showed that teachers' perception is driven by the social class of students in determining student outcomes. He provides an account of teachers' initial negative beliefs of students from disadvantaged social classes corresponding with lower levels of student achievement. Moreover, some experts found that the test score gap between white and African American students is perpetuated by teacher perceptions (Ferguson, 2003 and others).

The role of self-fulfilling prophecy in taking actions to improve student outcomes has been visited by many authors. If teachers expect students to succeed and student learning teachers' expectations increase effort to succeed then teachers' expectation is creating self-fulfilling prophecies (Brophy, 1983; Jussim and Harber, 2005; Jussim and Eccles, 1992). Ehrenberg and Brewer (1995) found that instruction from an African-American teacher was associated with higher gain in scores for African-American high school students. Gershenson et al. (2018) mention race or culture relevant pedagogy as another channel through which race-matching may operate.

Existing literature shows that teachers' motivational inputs can affect not only class-specific student achievement but also can bolster students' motivation to achieve future academic goals (Spera and Wentzel, 2003). Teachers evaluate their students' works and can make them aware of their mistakes, give valuable suggestions, praise them for their successes; these inputs can reinforce an average student's confidence in a class and motivate them to work harder in class in the future (Turner et al., 2002; Urdan and Schoenfelder, 2010). The challenges students face in a class regarding tasks such as disruption in class, peer abilities like teachers' appreciation and encouragement can mollify the effects of the negative elements (Bandura, 1986; Dweck and Leggett, 1988). Outside instructional motivation, designing classes to increase student autonomy can be effective in increasing student performance (Stefanou et al., 2013). Although teachers' subjective motivational inputs are assumed to share a direct relation with student outcomes, education literature pays little attention to identifying specific channels of the interplay between them. The self-fulfilling prophecy is one channel that economic theory assumes to be responsible for the interaction between subjective inputs and student outcomes[5]. The channel

[5]In recent economic literature, subjective measure of student and teacher quality has been widely used. Some of the studies that estimate the effect of subjective evaluation scores of teachers on student

of the Role-model effect discussed in section 1 functions in the same way as creating self-fulfilling prophecies. More recently, evidence of the golem effect or the opposite of the Pygmalion effect has been found in Dee (2004, 2005), Fairlie et al. (2014), Gershenson et al. (2016), and Sansone (2017). The golem effect in case of demographic mismatch is associated with negative student outcomes or negative effects on student outcomes.

3 The Data

In this section, I discuss the data used in the study and discuss descriptive statistics based on the data. Section 3.1 discusses the data source and Section 3.2 discusses the statistical description of some of the key variables that are later used in the empirical analysis of the paper.

3.1 High School Longitudinal Study of 2009 (HSLS:09)

The data I use in this paper comes from the High School Longitudinal Study of 2009 (HSLS:2009)[6], which is a nationally representative survey sponsored by the U.S. Department of Education. This survey tracks a single cohort of U.S. high school students drawn from the universe of 9th graders in the fall of 2009[7]. The base year survey was conducted in 2009. In this paper, I use the data from the base year survey. I use the restricted-access data files from the Institute of Education Sciences (IES) that cover teacher survey, high-school administrator or principal survey, high-school counselor survey besides the student survey. The 9th-grade science and math GPA has been calculated from the restricted-access student-course file in HSLS:09. The student-course files contain all course information of the respondent students[8].

achievement use evaluations from professional agencies that specialize in training, mentoring and rating teachers' effectiveness (Rockoff and Speroni, 2010, ; Kane et al., 2011; Bacher-Hicks et al., 2017; Blazar, 2015; Kane and Staiger, 2008; Kane and Staiger, 2012).

[6]Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

[7]Schools and Students are the data sampling units, not parents; the response rates in terms of schools and students.

[8]The high schools participated in HSLS:09 provided the course information directly. I calculated the grade points for 9th grade math and science using the same categorization used to create high school math and science GPA in the publicly-available "student" files (X3TGPAMTH and X3TGPASCI, respectively). The "school course" files supplied by IES directly provide additional categorizations such as "pass" and "fail" as course outcomes. Using the same categorization, I allocate the median grade of the course to "pass" and the failing grade (F) to "fail" category. Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

The HSLS:2009 contains detailed information about students' academic achievement, socioeconomic background, demographic characteristics, educational attainment of parents, parents' occupation, etc. The survey collects data on home life and background information of the parents, administrators (principals), and high school counselors. From the sampled schools, math and science teachers also participated in the survey through a separate questionnaire[9]. HSLS:09 employed a stratified, two-stage random sample design with primary sampling units defined as schools in the first stage, and students were randomly selected from the sampled schools in the second stage. HSLS:09 defined its target population as students from 944 public schools, including public charter schools, and private schools in the 50 states and the District of Colombia that provided instruction to 9th grade students in Fall 2009. Ingles et al. (2011) and Duprey et al. (2018) provide detailed information for school inclusion and exclusion rules. They also discuss further details about administering the survey and sample selection. A sample of 25,210 study-eligible students was retained in the published data.

I create an unbalanced panel of distinct 9th-grade students, where each student has two observations for math and science. 82% students have two observations, 18% students have only one observations. This imbalance is primarily due to non-response from their teachers. The total number of student-teacher observations in the unbalanced panel is 16,730[10]. All unweighted sample sizes are rounded off to tens as per IES regulation. HSLS:09 asked students to give information about teachers when they conducted the student survey. Teachers were contacted after the student survey and asked if they wanted to participate in the survey. The student survey file has a cross-sectional layout where the teacher responses and principal responses (columns) are identified with student and school IDs in each row. HSLS does not include unique IDs for teachers, but grouping together some teacher responses for each school allowed creating unique teacher ID variables[11]. These teacher response

[9]The published data links teachers, parents, administrators, high school counselors to students.

[10]The math teacher survey data had over 6,430, science teacher survey data had 7,940 and principal survey data had about 1,500 missing responses. Parent survey in HSLS:09 has additional 6,500 missing observations. Parent survey variables are not necessary to be included in the analytical sample as they are invariant within a student.

[11]Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

variables are race, sex, highest degree earned, year of highest degree earned, teachers' highest degree in STEM code, whether math teacher's BA/BS degree awarded by the education department, the type of teaching certificate currently held, years of experience in teaching students in 9th to 12th grade, and lastly, whether the teacher was collecting from teacher retirement system/401(k)/403(b). For my analysis, it is not necessary to find unique teachers, but later I use it for a sensitivity analysis. School *times* Unique teacher fixed effects controls for the varied effect within each school the different math/science teacher may have on student achievement.

3.2 Key Instruments of Analytical Sample

The main analysis in this paper uses a panel of 9th-grade students and math and science teacher response variables. The main outcome variable, 9th-grade GPA, varies within math and science courses and so do the explanatory variables. Every observation is unique at the Student$\times$course level. Student ID is the panel variable. Table 1 presents statistical description of the key outcome variables and some other measures of student achievement. In Table 1, Columns 3, 4, 6 refer to mean-difference t-tests between Column 2 (White Students) and 3 (Non-white Students), 2 (White students) and 4 (Black students), and 5 (Male Students) and 6 (Female students), respectively. The stars signify the p-values of the t-test statistic[12]. Table 1, shows that mean STEM GPA (9th grade) is significantly different between White and non-White students, and White and Black students[13]. Mean GPA is higher than that of non-White and Black students. On the other hand, the Mean GPA for female students is significantly higher than that of male students; a fact that indicates the gender gap reversal in education (Hershbein (2013)). The This pattern of demographic differences in achievement is true for other student outcomes such as math/science AP credits in high school, high school math/science GPA, and self-efficacy in math/science[14]. The minority students - non-White and Black - on

[12]The stars or asterisks ***, **, * refer to ***p<0.01, **p<0.05, and *p<0.1 of the statistic of the mean-difference t-tests.

[13]GPA is measured on a standard 4-point scale.

[14]Self-efficacy is a standardized variable that presents the scale of the student's efficacy in math or science. This variable is provided in HSLS:09 and is created through principal components factor analysis. The inputs to this scale were variables that rated (1) how prepared the student feels about the math/science tests, (2) how well the student understands the math/science textbooks, (3) how confident the student is about his/her skills in math/science, and (4) how confident the student is that she/he would excel the

average face more teachers of a different race, a smaller number of certified teachers, slightly smaller number of teachers with STEM degrees.

Table 2 presents means of the same variables as table 1 for samples split by teacher types. Black and non-white teachers face smaller average 9th-grade GPAs than the students they teach. This is consistent with the economic literature which shows that white teachers are more likely to teach in higher-performing schools and that Black teachers tend to shift to schools with a greater percentage of Black students(Hanushek et al., 2004)[15]. Female teachers on average have a smaller number of STEM degrees and are less experienced as Table 2 shows. However, female students have higher levels of achievement according to Table 1, except for self-efficacy. This pattern of higher levels of female achievement compared to the males' is consistent with the reversal of the gender gap in education(Hershbein, 2013). Family income and parents' education are two major drivers of student achievement. As these two factors are invariant within a student, I do not include them in the panel analysis[16]. But it is interesting to see how these two variables vary across races. The mean (annual) family income corresponding to the White student sample is approximately 20 thousand dollars higher than that in the non-White sample and even higher than that in the Black sample (Table A.2).

Overall, White teacher and White student sample means resemble the total sample means of the variables. White students have stand out as having higher levels of average student achievement over non-White and Black students which education research often shows (Arcidiacono et al., 2012; Fryer Jr. and Katz, 2013 and others).

math/science course on Likert scale.

[15] Black teachers on average are assigned to a different-sex and different-race students in a sample of multiple schools in different school districts when school and higher-order heterogeneity are controlled for (Hanushek et al., 2004).

[16] Moreover, variables from parent survey in HSLS:09 suffers from a high degree of missing responses. The variables for family income and parents' education are both formulated from various responses in the parent survey.

4 Identification

The main model for this paper has the same structure of the "Education Production Function" introduced in Krueger (1999)[17]. Class size is input in his education production function. The basic structure of the education production function reflects the relationship between outcome Y_{ij} and inputs such as indicators of demographic mismatch, achievement controls, etc.

$$Y_{ij} = f(D_{ij}, X, \epsilon_{ij}) \text{ where } j \in \{S, M\} \tag{1}$$

In the equation 1, S stands for Science and M for Math. The above equation shows that student achievement y_i is a function of educational inputs for student i in a sample of N students. $\epsilon_i j$ is an error vector with a normal distribution. D_{ij} is the vector of the **demographic mismatch indicators**.

Formally, the STEM achievement is modeled as equation 2. This is a panel model where the outcome and explanatory variables vary across the courses and within each student. Student-specific effects - both unobserved and observed, and invariant *within* each student - on the outcome are thus controlled for. This identification strategy has also been implemented by Dee (2005), Fairlie et al. (2014), Gershenson et al. (2016), and Sansone (2017), and others.

$$Y_{ij} = \alpha_j + \gamma_i + \mathbf{X}'_{ij}\beta + \Omega D_{ij} + \epsilon_{ij} \;\; \forall \; j \in \{S, M\} \tag{2}$$

In equation 2, α is a subject fixed effect (either math or science) that controls for systemic differences in these two courses. γ is a student fixed effect that controls for unobserved student characteristics. In a pooled model without γ, or where data is available for only one course, γ would be included in the error term, making the estimate of Ω biased. Math and science skills overlap in certain ways, e.g. mathematical computation is both necessary in geometry, algebra, and physics. Because of this overlap in skills, γ captures a greater degree of unobserved effects

[17] Education production function models student achievement (SAT score of the students in Tennessee STAR school) – on independent variables such as student characteristics, class size, etc. His prior was smaller class size would increase student achievement.

than it would in a model where achievement varies between reading or writing, and math courses (e.g. Sansone, 2017; Papageorge et al., 2019). Ω captures the effect of demographic mismatch (race-mismatch and gender-mismatch) compared to race-matching and gender-matching between the teacher and the student.

X is a vector of observed teacher characteristics such as experience, whether he/she has a master's or higher degree, and whether the highest degree was in STEM. In addition, science and math teachers' responses to whether disruption and students dropping out are major problems are added as control variables. A wide range of student-level unobserved and observed factors related to study efforts and past achievement are controlled for when student fixed effect is included in the model. To capture the differential effect of efforts in math and science courses I additionally include daily homework hours and daily homework hours squared as control variables. D_{ij} is a demographic mismatch indicator.

Following Dee (2004) and Gershenson et al. (2016), I further formulate D_{ij} as a vector of four mutually exclusive categories teacher-student demographic mismatch: same race and same sex, different sex and same race, different race and same-sex, and different sex and different race. The best strategy for categorizing demographic mismatch is using the same-race-same-sex as the reference category as the expected signs of all the other categories would be negative. If other-race-other-sex were the reference category, then "same race but other sex" could have either sign because same race is an improvement on the reference category but other sex is not. A negative sign could reflect the "other" neutralizing the "same". A positive sign of would reflect the opposite. Such a categorization would definitely complicate our analysis. The effect of demographic mismatch or matching between teachers and students is likely to be more prominent for minority students. Among the main minority groups - Black, Hispanic and Asian, Black students and Hispanic students have a relatively higher percentage of racial match with the teachers - approximately 5%. However, only 1.5% of Asian students see the least percentage of racial matches with their teachers. In this paper, therefore, the main focus is on the demographic mismatch impact for Black students.

For Asian students, this reference category "same-race-same-sex" is smaller in terms of the number of observations compared to other minority groups. This issue leads to the problem of large coefficients of the D_{ij} variables because sparsely distributed observations may drive the results. The minority students are more likely to experience role-model effect or similar effects due to the presence of the same race teacher (Dee, 2004). Across minority races, this effect can vary, making it heterogeneous. Thus, I look at split the analytical samples into demographic categories to see the differences in the effect of demographic mismatch. The main outcome variable is z-scored. As the main explanatory variables - D_{ij} - cause differences in the average position of the students in their class. As the model has a within-student panel structure, the estimated effect of D_{ij} would be driven by the variations in demographic mismatch *between* math and science courses; on the average, a demographic mismatch in math or science class would likely cause a lower position in that class compared to the other class where student sees demographic "match" with the teacher. This setup assumes that the effect of demographic mismatch is the same for math and science.

5 Sorting Tests

HSLS:09 randomly selected students for the survey but the teachers were not randomly assigned to them. If student achievement is a factor in racial matching between teachers and students, then it poses a threat to the validity of the student fixed effect estimates. For example, if low-performing minority students are assigned to classes that a minority teacher teaches, then the estimated effect of a demographic match between the minority teacher and the student could be biased downward. On the other hand, if high-performing students are assigned to a minority teacher, then the effect of demographic mismatch on student outcomes could be biased upward. Therefore, I test if evidence of sorting exists for this sample using techniques implemented in Fairlie et al. (2014), Gershenson et al. (2016), Lusher et al. (2018), and Oliver et al. (2021).

The first test I conduct is similar to what Lusher et al. (2018) used. His data covers information on teaching assistants and students in a wide range of classes. He collapses the dataset at class level and then tests if the fraction of minority students is correlated to the fraction of TAs that teach on class-level or section-level. The data I use only covers two classes (math and science) and the majority of schools (67%) in the analytical sample has at most 4 teachers taking the HSLS survey. Therefore, I simplify Lusher et al. (2018)'s test with the following model.

$$\begin{aligned} NWT_{ijk} = \beta NWS_{ijk} + \eta \mathbf{Achievement}_{jk} + \Omega NWS_{ijk} \times \mathbf{Achievement}_{jk} + \\ School_k \times Course_j + \epsilon_{ijk} \quad (3) \\ \forall\ j \in \{S, M\}\ \ \&\ \ k = 1, 2, 3..., 944 \end{aligned}$$

NWT_{ijk} is a binary indicator of a non-White teacher who teaches course j in school k to student i. NWS_{ijk} is a binary indicator of a non-White student with the same subscripts. For $\mathbf{Achievement}_{jk}$, I use the teacher-specific average theta score and standardized grade in the highest math/science course the student took in 8th grade[18]. $School_k \times Course_j$ fixed effects should capture systemic differences in school-specific math and science courses. The main coefficients of interest are β and Ω. If they are statistically significant then the sorting test would say that students' racial minority identity is associated with that of the teacher through the student-teacher assignment. Identification of Ω requires there to be multiple teachers for a given school-course. Only 93% of the school-courses have more than 2 unique teachers[19].

The next sorting test I employ is from Gershenson et al. (2016). They use data from the Education Longitudinal Study of 2002 (ELS:2002). This sorting test is similar to Lusher et al. (2018), in that the average teacher-specific achievement is modeled on a interaction term of two binary variables reflecting minority status of teacher and student. This test was originally developed by Fairlie et al. (2014). Their test checks for differential sorting on observable characteristics of students. They argued that if there is no sorting based on the observed characteristics, and

[18] The mathematical theta score is the the score on a algebra test that each respondent took. The score has been z-scored within the sample. More on mathematical theta score can be found in Ingels et al. (2011)

[19] *High School Longitudinal Study of 2009 (HSLS:09/16), Base-year Survey* (2009) does not provide unique teacher ids. By finding unique combinations of the teacher responses to a range of questions (sex, race, years of teaching experience, highest degree, bachelor's degree major, certification type, pension indicator) I create teacher IDs.

mean teacher-specific measure of student quality is not correlated with the error in Equation 2, then differential sorting would not threaten the validity of the student fixed-effects estimates in Equation 2.

$$\begin{aligned}\overline{Achievement}_{jk} = \beta NWS_{ijk} + \eta NWT_{ijk} + \Omega NWS_{ijk} \times NWT_{ijk} + \\ School_k \times Course_j + \epsilon_{ijk} \\ \forall\ j \in \{S, M\}\ \ \&\ \ k = 1, 2, 3..., 944\end{aligned} \tag{4}$$

This test sees if the average quality of the students a minority teacher teaches is associated with the match between minority teachers and students. For example, if the average school assigns low-performing minority non-White students to minority teachers then Ω would have a negative sign and be statistically significant. On the other hand, if the average school assigns high-performing minority non-White students to minority teachers then the estimated Ω would be positive. Essentially, Ω estimates how the mean difference between White and non-White student characteristics varies between White and non-White teachers. If estimated Ω is statistically indistinguishable from zero, it would suggest that there is no evidence of race-specific sorting based on observable student characteristics. In that case, sorting based on unobserved characteristics of students is unlikely.

6 Results

In this section, the empirical analysis is presented. The subsections are arranged in terms of the complexities of the estimated models. The sorting tests and the other tests are discussed in section 6.1. From section 6.1 to section 6.4, I discuss results mostly from pooled regressions. The main model for analysis has a panel structure, as Equation 2 shows. Section 6.5 to section 6.7 discuss the baseline results, robustness tests, and the effect of demographic mismatch on alternative student outcomes. These sections present results that mainly used the panel model with student fixed-effects.

6.1 Sorting Test Results

The first sorting results are shown in Table 3. This table shows the estimates of Equation 3. The indicator of non-white students does not appear to have and statistically significant association with the Non-white teacher indicator (NWT). The interaction terms - Teacher-specific average theta score (MEANTHETA) $\times$ Non-white Student (=1) and Teacher-specific average z-scored grade of highest math/science course in 8th grade (ZG8) $\times$ Non-white Student (=1) - have coefficient estimates statistically indistinguishable from zero.

The second round of sorting test results comes from estimating Equation 4. The main coefficient of interest is Ω in this equation. The estimated results are in Table 4. The interaction term - Non-White Student $\times$ Non-White Teacher does not appear to have any significant impact on the average MEANTHETA or ZG8. Columns 1 and 2 results in Table 4 are estimated with the inclusion of School$\times$Course fixed effects and columns 3 and 4 only with course fixed effects. The test results suggest no differential sorting on observed characteristics by student race. Therefore, differential sorting on unobserved factors is unlikely to threaten the validity of the identification of student fixed effects. In addition to these results, Table A.5 provides test results of sorting on family income and respondent parents' education. PAR1COL is a binary variable, coded 1 if the respondent parent or parent 1 has a college degree. In all columns in Table A.5, the interaction term between non-White students and non-White teachers does not produce any significant impact on the observed characteristics - parents' education and family income.

6.2 Falsification Tests

Table 4 results also can be interpreted as results of falsification tests. Rothstein (2010) tested whether present achievement is connected to the teachers to whom students are assigned in future in a value-added model. The future teachers' effect on present achievement if is found to be statistically significant, that would mean the correlation between the input and the outcome is spurious. After Rothstein (2010), many others applied different versions of this test with different education production models. Table 4 shows that unique teacher-specific mean ZG8, the z-scored 8thgrade math and science achievement is not impacted by Non-white student × Non-white Teacher in a significant way. MEANTHETA is a algebra test score z-scored within the whole sample. It may be related to 9th-grade teacher's input, as the baseline survey was conducted in Fall 2009, when the students were already in 9th grade. If a student, on the other hand, took the test very early in Fall 2009, then the connection between Non-white student × Non-white Teacher would be very weak. Table 4 shows that Non-white student × Non-white Teacher does not impact unique teacher-specific MEANTHETA score in a significant way either.

Columns 2 and 4 in Table 4 show that Non-white student × Non-white Teacher does not have a significant impact on 8th grade math/science GPA of the participant students. I also do an additional falsification test that is based on the main model, Equation 2.

$$Y_{ij,t-1} = \alpha_{j,t} + \gamma_{i,t} + \mathbf{X}'_{ij,t}\beta + \Omega D_{ij,t} + \epsilon_{ij,t} \;\; \forall \; j \in \{S, M\} \tag{5}$$

The falsification test would measure the effect of $D_{ij,t}$ on past achievement - $Y_{ij,t-1}$ - in the context of the current analysis. Since any teacher-survey information is not available for any period later than the Fall 2009, it is beyond the scope of this paper to estimate the effect of $D_{ij,t+1}$ on $Y_{ij,t}$[20]. Table A.8, shows that there is no significant impact of $D_{ij,t}$ on $Y_{ij,t-1}$. To ensure, that students were not matched with their current math and science teachers in the 8th grade, I only consider the samples of students who were attending schools in Fall 2009 that did not include 8th grade. None of the coefficient estimates of the demographic mismatch categories are statistically significant, which means it is not likely that the effect of demographic mismatch in Fall 2009 is spuriously connected to achievement through some unknown factor that also drives past achievement.

[20]The effect of $D_{ij,t}$ on $Y_{ij,t+1}$ could be estimated to see if the relationship between demographic mismatch and achievement was spurious, but it is possible that demographic mismatch would also continue in the next grade. In that case, an additional measure to separate future demographic mismatch, $D_{ij,t+1}$, from the $D_{ij,t}$ would be necessary. However, HSLS:09 does not include a teacher survey after baseline wave.

6.3 Selection Bias

Sorting tests have established that demographic match or mismatch in the analytical sample is not related to average teacher-specific student achievement. In the survey sample, not all students take math and science in 9th grade. If selection into math/science courses is merit-based, then the impact of demographic mismatch may be biased downward - more talented students who are selected into STEM courses would find it easier to battle the negative impact of the mismatch. The panel structure weeds out the unobserved and observed invariant factors that might have led to selection in the STEM classes. Moreover, in 9th-grade students take foundational or basic STEM classes which may be necessary for high school completion. This set-up would limit the possibility of self-selection into different classes. I conduct an additional test to see if selection into Fall 2009 math/science classes was merit-based.

The selection test utilizes the panel structure of the data and estimates a student fixed-effects model, where the main independent variable is standardized grade in the highest-level math/science course in 8th grade and the outcome is whether math was taken in Fall 2009. The result of this test is in Table 5. The coefficient estimate for ZG8 is not statistically significant which shows there is little scope of merit-based selection into fall 2009 math and science classes. I use all the observations for this test that were available. The sample used for this test is considerably larger as teacher and principal survey variables - which are major sources of missing values in HSLS:09 - are absent here. Not all students take both math and science courses. Some students take math but not science and some take science but not math. Therefore, *within* students, the course-selection indicator would vary. In 9th grade, students take mainly courses that are required for graduation - algebra, integrated math, geometry, etc. So, in 9th grade achievement-based course selection were much less likely.

6.4 Teachers' Race and Student Achievement

"Pre-baseline" results from a pooled regression analysis allow the readers to understand how the effect of demographic matching changes depending on what model we fit over the sample. The main model of interest in this paper uses a panel data structure. Such a model controls for the factors invariant to the panel variable. A "pre-baseline" model offers more flexibility to observe race-related patterns. In this paper, the main outcome of interest is 9th-grade GPA varying between science and math. 9th-grade GPAs are z-scored at the school level. Table 6 provides results from a pooled regression. Results in each column in this table differ only in race categories of the teachers. When the whole sample is considered, Black and Hispanic teachers have negative and White and Asian teachers have positive impacts on STEM achievement. These "raw" results from the pooled regression do not account for the past achievement, the various observed and unobserved drivers of student achievement, the effect of school, etc. Therefore, the race indicators are most likely picking up average school quality in addition to their own effect.

In Table 7 we see a different picture after splitting the student sample by race. In column 1, I include all teacher race indicators keeping White teachers as the reference group. None of the race categories appear to generate any significant impact on 9th-grade GPA as shown in the first column. In columns 2, 3, and 4, I fit the pooled model on the samples of Black students, Hispanic students, and Asian students respectively. The results in Table 7 tells a different story than Table 6's. The impact of Black students being taught by a Black teacher is positive and statistically significant. The same is true for the Hispanic student sample. In both Tables, 6 and 7, the pooled model only considers a few control variables. Family income and parents' educational achievement are not considered as controls. Table A.3 includes region and locale control variables besides family income and one parents' educational indicator. For all student samples, family income and parent 1's education (whether he/she has a college degree) produce positive impacts on 9th grade standardized GPA. In column 2, Table A.4, the estimated effect of having a Black teacher on Black students is statistically significant. In Columns, 3, 4, and 5, the effect of having the same-race teacher is not statistically significant. In Table A.4, the teachers' race indicators are formulated differently; three variables presented are - whether student faces teacher with a different race, whether she/he faces teacher of the other sex or both. The whole sample, including the Black sample of students experience negative impacts from having teachers of a different racial identity.

6.5 Baseline Results

Estimates of Equation 2 are presented in Table 8. Having a different-sex teacher appears to have a statistically significant negative impact on Z-scored 9th grade GPA in math/science courses (ZGPA9). Its effect is statistically indistinguishable from zero for the Hispanic and female samples. The Asian sample sees a statistically significant impact of all three demographic mismatch categories on ZGPA9. Further stratification shows that only Black female students face the statistically significant negative impact of having an other-sex and both other-race and other sex teacher (Table 9). The intersection of race and gender by splitting a relatively small sample of Asian students has possibly weakened the effect of demographic mismatch on them (Columns 8 and 9).

Moreover, Asian students in the sample are rarely matched with Asian teachers. The results in Column 8 could be driven by very few instances of racial matching, which could make them exceedingly large. The *within* student panel does not require the inclusion of variables that are invariant within students. The baseline model includes teacher and school-related control variables that vary between the science and math courses or science and math teachers. HSLS:09 provides a wide variety of teacher-survey instruments. I include a parsimonious selection of teacher-survey instruments as controls because teacher quality and school quality are not likely to vary widely even though at least two different teachers (math and science) are reporting them. Teacher instruments are *whether the highest degree major was in a STEM field, years of experience in teaching high school students, percentage of teachers from Spring 2009 returning to teach in Fall 2009, highest degree earned by the teacher, whether advanced math courses are assigned to all or most math teachers, whether students dropping out is a problem.* Student effort put into studying math and science might not be the same. There is no explicit measure of effort levels varying between the two courses. I include hours spent doing math/science homework and homework^2 per week as two additional control variables in the baseline equation.

The rest of this section is dedicated to discussing robustness checks of the baseline results in Tables 9 and 10, and investigating potential channels through which demographic mismatch between teachers and students impacts student achievement. The robustness checks include various measures of student achievement that could be affecting 9th-grade math/science GPA via demographic mismatch. I extend the baseline results by including past achievement as a control variable, highest-level 8th-grade math, and science course indicators and 9th-grade course indicators in Tables 10, 14, and 12. These extensions also serve the purpose of a sensitivity analysis. First, Table 10 shows that including past grade ZG8 as a control variable leaves the coefficients in Table 9 unchanged. Second, in Table 14, the baseline regression has been extended by both ZG8 and the indicator of the highest STEM course of 8th grade. In Table 14 too, the results do not change from the baseline. In addition, the other-race and same-sex categories for Black female students are statistically significant.

I develop alternative regression specifications by adding different types of fixed effects and control variables to equation 2. I use the more stratified samples like in table 9 whenever possible for the robustness checks. However, splitting the sample by both races leaves too few observations for some of the robustness checks. Further, when interaction terms are applied, some demographic categories - like those in Table 9 - are not identifiable due to a lack of variation in their values. In those cases, I use the same sample division like that in Table 10.

I add 9th-grade math and science course indicators as controls in Table 12. This extension produces demographic mismatch estimates for Black females that are statistically indistinguishable from zero. However, same-sex and other-race for Asian female students are statistically significant. In another alternative specification, I add the unique teacher fixed effects besides course fixed effects to the baseline model in Equation 2 and estimate the effect of demographic mismatch. The coefficient estimates in Column 6 in Table 13 for Black female students are close to the baseline in Table 9.

HSLS:09 respondent students mention factors such as parental encouragement, teachers' encouragement, personal enjoyment, career development, and love for the challenge as reasons for taking STEM courses. The statistical description of these factors is provided in Table A.1. In Table 15, the baseline model is extended with interaction terms between D_{ij} and teachers' encouragement. Teachers' encouragement - TEACHER in Table 15 - is coded 1 when students report teachers' encouragement as a reason for taking math in 9th grade. In the lower panel, marginal effects of a teacher at levels of D_{ij} are shown. In columns 1 and 5 of Table 15, for Hispanic and Asian students, teachers' encouragement generates positive impacts on achievement, neutralizing the negative impacts of demographic mismatch. On the other hand, in columns 1 and 5, the negative effect of a sex-mismatch between teachers and students persists, and it outperforms the effect of teachers' encouragement. Overall, Table 15 shows that the negative effect of demographic mismatch can outweigh the positive motivational influence of teachers.

The student fixed effects in equation 2 are not only useful for controlling for the student-specific observed and unobserved factors but also for the effect of the average quality of the teachers, their beliefs, and mindsets. HSLS:09 provides a list of science and math teachers' beliefs and attitudes[21]. This list of variables allows us to control for attitudinal effects from teachers affecting achievement through demographic mismatch. To elaborate with an example, the statistical description in Tables 1 and 2 showed that Black students[22], on average, see lower levels of student achievement. If White teachers in Black-majority schools are of low quality where quality is heterogeneous across courses and schools, and an aspect of being low-quality is being skeptical about student success, then demographic mismatch would be picking up the negative effect of teacher beliefs and attributes that are at odds with promoting student success. To check if that is the case, I include some teacher belief indicators that might have a negative impact on student outcomes in the baseline model (equation 2) and provide the estimated results in Table A.7. The teacher belief variables are described in Table A.2. Compared to the baseline results in Table 9, the coefficient estimates of demographic mismatch in Table A.7 are not different by a considerable margin.

[21] Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

[22] In a future draft, a separate table will be created to show the means specific to math and science classes.

6.6 School Characteristics & Demographic Mismatch

In this section, I explore whether certain school characteristics can cause the impact of demographic mismatch to be heterogeneous with a demographic group. HSLS:09 interviewed the school administrators and collected information on a wide range of school characteristics[23]. The importance of examining the relevance of these factors lies in the possibility that teachers' qualities can increase or decrease with school-related variables; these variables can also exhibit a pattern of correlation with the demographic composition of high school faculty. Some of the school characteristics I visit in this section are (1) percentage of 9th-grade repeaters, (2) percentage capacity to which the school is filled, (3) percentage of the student body is Black, (4) percentage of students enrolled in free lunch, and lastly, (5) whether bullying is a problem in the school. All the regression results discussed in this section are estimated using the same specification as that of the baseline results - Tables 8 and 9. Table A.1 provides sample means of these characteristics. Higher percentages in the first two variables may correspond to the lower quality of a school[24]. The higher percentage of students enrolled in free lunch programs may also indicate that more students face food insecurity, which is associated with poverty. However, smaller samples created to examine these school characteristics mean a lack of variation in the key variables of interest. As I previously mentioned, minority students in the sample rarely are assigned to minority teachers of the same ethnic/racial background. This causes the results to be driven by few instances of a demographic match. This is the reason why for the sample of Asian students the coefficient estimates are very large; close to 1 standard deviation. For these two reasons, most of the results discussed in this section reflect that in smaller samples the key variables have weaker effects.

The median percentage of students repeating 9th grade is 1%. This is the first school characteristic I discuss here. Figure 1 shows that the Black and Asian students face the statistically significant negative impact of demographic mismatch in schools with 1% or more 9th grade repeaters. All three categories of demographic

[23] Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

[24] One caveat is that schools not running on full-capacity does not guarantee higher institutional quality. If the enrollment rate low and a school cannot fill a substantial number of seats, then that would also indicate low capacity. Further development of this research will consider this issue.

mismatch have similar impact magnitudes for Black students. "Other Sex" category for Asian students dropped out due to lack of variation. Interestingly, in schools where less than 1% students repeat 9th grade, the estimated impacts of demographic mismatch are statistically indistinguishable from zero for all races. So, Panel (a) and (b) show lower-quality schools with more than 1% 9th grade repeaters may see a negative impact of demographic mismatch.

Next, I discuss school capacity saturation as a source of heterogeneous impact of demographic mismatch. Figure 2 shows that the impact of "other sex" teachers in the classroom is almost the same for Black students in both types of schools - the schools filled 95% of its capacity and the schools filled 95% or less than its capacity. Besides Black students, the results for Asian students are statistically significant only for the set of schools that are filled 95% or less.

Further, I look at schools where Black students are the majority. Black females experience statistically significant effect of demographic mismatch in only Black majority schools (Figure 3). The effect of demographic mismatch is not statistically significant in Black-minority schools. This same exercise could not be implemented for other races, as there are less than 10 Hispanic and Asian majority schools in the analytical sample.

The next two figures, Figure 4 and 5, do not provide noteworthy insights into the heterogeneity of the impact of demographic mismatch because the majority of the estimates in these two figures are statistically indistinguishable from zero. The statistically significant result in Panel (a), Figure 4 shows that the impact of demographic mismatch for Asian students is approximately 0.8. On the other hand, in and (b) - schools where less than 50% students get free lunch, the impact of "other sex" is smaller in magnitude than that of Asian students. Figure 5, Panel (a) shows that in schools where bullying is a problem, Asian students see a loss of -1.2 standard deviations when assigned to an "other sex" teacher. On the other hand, Black students experience a loss of 0.53 standard deviations for the same type of demographic mismatch in schools where bullying is not a problem.

6.7 Alternative Outcomes

In addition, to 9th grade STEM GPA, the main outcome variable of this paper, I look at the effect of demographic mismatch on a few other student achievement measures. Table 16 and 17 provide the estimated effect of demographic mismatch on total math/science AP credits in high school (APC) and high-school math/science GPA respectively (HGPA). High school GPA in math and science are calculated by collecting all math and science course grade points from the 9th grade. A statistically significant impact of demographic mismatch on AP credits and High school GPA would mean that demographic mismatch not only negatively affects current course outcomes, but it can worsen longer-term educational outcomes too[25].

Table 16 shows demographic mismatch lowers AP credits for Black males and Asian females. On the other hand, demographic mismatch in Table 17 negatively affects a wider range of samples, White females, Hispanic females, Hispanic males, Black males, and Asian females. Compared to the baseline results in Tables 8 and 9, the results in Tables 16 and 17 exhibit that demographic mismatch have longer-term negative impact lasting throughout the high school lives of some demographic groups. In comparison, the baseline results did not show any negative impact for Hispanic female students, Hispanic male students, and Black male students.

As HGPA and APC are later outcomes than 9th-grade GPA (ZGPA9), they may capture demographic mismatch exacerbated by mismatch in multiple grades or classes. On the other hand, estimates lacking statistical significance may be driven by the fact that demographic mismatch led to self-selection (or outcome-driven school-authorized) out of elective math and science courses and thus weakened the link between end-of-high-school outcomes such as HGPA and APC[26]. Further, This allows the possibility that the students would have some opportunity to match with teachers they like, and thus possibly deflates the effect of initial demographic mismatch in 9th grade. These findings about APC and HGPA, mainly hint at the longevity of the negative effect of demographic mismatch but they may also

[25] HGPA is standardized at school level. APC contains only positive numbers and it can be compared across school, so it is not transformed or standardized.

[26] One caveat is that the most schools require a minimum number of math and science classes to graduate, so the scope for opting out is mostly at the three- or four-course margin.

be biased. One cannot identify the proper channel through which the effect is conducted. A model that contains outcomes for every grade identifies could be used to address the question of longevity of demographic-mismatch-related loss for students.

In addition to these results, Table A.6 shows the impact oof demographic mismatch on 9th-grade math/science outcome[27]. A lot of authors examine self-efficacy in math, science, and other courses as important inputs to education production(see Chen and Usher, 2013). As multiple student-specific opinions of own competence are combined together, the exact source of a student's efficacy that is connected to demographic mismatch cannot be definitively outlined. Moreover, self-efficacy is constructed using subjective inputs from students. The average minority student assigned to a "other-race-and-other-sex" teacher may start feeling less confident about the factors that make up self-efficacy. I show the impact of demographic mismatch on self-efficacy in Table A.6. The results show that the effect of demographic mismatch on self-efficacy is very strong for Black males and Asian males.

7 Conclusion

In this paper, I present a model to estimate the impact of demographic mismatch on 9th-grade math/science course achievements. The importance of studying the effect of demographic mismatch on math/science outcomes lies in the fact that high-school math and science courses are foundations of future tertiary STEM education and that demographic test score gaps are transmitted to post-secondary levels. I use a unique nationally representative survey data of 9th-grade high school students in the U.S. that contain demographic and career information of math and science teachers. The data allows me to analyze within-student variations of achievement and employ student fixed-effects to identify the effect of demographic mismatch with their teachers. I do not find any evidence of differential sorting or merit-based

[27] Self-efficacy is a standardized variable that presents the scale of the student's efficacy in math or science. This variable is provided in HSLS:09 and is created through principal components factor analysis. The inputs to this scale were variables that rated how prepared the student feels about the math/science tests, how well the student understands the math/science textbooks, how confident the student is about his/her skills in math/science, and how confident the student is that she/he would excel the math/science course. The answers to these questions were recorded on Likert scale.

math/science course selection.

The identification strategy employed here was introduced by Dee (2005) in addressing the question of "role-model effect". Many others, including Fairlie et al. (2014), Gershenson et al. (2016), Sansone (2017), and Gershenson et al. (2018) used this identification strategy. The analysis shows consistent results that having a different-sex teacher generates a negative impact on STEM achievement. Having a different-sex and different-race race generates comparatively weaker effects on achievement. As Black teachers and Black students are both minority groups, and the survey did not oversample Black teachers, occurrences of race-matching is are few. For this reason, different-race and different-sex effect is found to be statistically less robust. Black female students see a significant loss in achievement (approximately 0.3 standard deviations) due to being assigned to a different-sex and different-race teacher in the 9th grade. Moreover, Asian students see the greater impact of having a different-race teacher, a different-sex teacher, and both. However, a demographic match for Asian students are rare, and the coefficient estimates may be exceedingly large because very few instances of mismatch showing very large impacts of over 1 standard deviation. Within every racial/ethnic group, neither males or females are minority, and every group sees almost the equal number of sex-match and sex-mismatch. That is why different-sex teacher effect is statistically more robust in the evidence provided in this paper.

Overall, the results complement the findings in Gershenson et al. (2016), Fairlie et al. (2014), and Sansone (2017). The systemic negative effect of demographic mismatch deserves more research, especially how demographic mismatch operates in the classroom. The channels through which demographic mismatch operates has been shed light on in this paper briefly. Using students' self-reported reasons for taking math and science in 9th grade such as teachers' encouragement, where course choice is very limited, I show that teachers' encouragement has some positive impact on some subsets of minority students, on the average, but its positive effect can be neutralized by the demographic mismatch. Therefore, demographic mismatch may just be operating through the channel of passive or direct demotivation. This channel is similar to what is known as the "Golem effect" in the literature.

The findings in this paper show that school quality and characteristics may influence the channel of demographic mismatch. I find that possible overcrowding in schools corresponds with a greater negative impact of a demographic mismatch for Black students. Other evidence in this paper suggests that a higher percentage of 9th-grade repeaters in a school may accentuate the negative effect of a demographic mismatch for Black students. Further research in this area is needed that could explain the connection between school characteristics and demographic mismatch.

References

Altonji, J. G. (1992). The effects of high school curriculum on education and labor market outcomes. (w4142). https://doi.org/http://dx.doi.org/10.3386/w4142

Arcidiacono, P., Aucejo, E. M., & Spenner, K. (2012). What happens after enrollment? an analysis of the time path of racial differences in gpa and major choice. *IZA Journal of Labor Economics*, *1*(1), 1–24.

Bacher-Hicks, A., Chin, M., Kane, T., & Staiger, D. (2017). An Evaluation of Bias in Three Measures of Teacher Quality: Value-Added, Classroom Observations, and Student Surveys. (w23478). https://doi.org/http://dx.doi.org/10.3386/w23478

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Bertrand, E., M.and Duflo. (2016). Handbook of field experiments. Cambridge: Abdul Latif Jameel Poverty Action Lab.

Blazar, D. (2015). Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review*, *48*(100), 16–29. https://doi.org/http://dx.doi.org/10.1016/j.econedurev.2015.05.005

Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of educational psychology*, *75*(5), 631–795. https://eric.ed.gov/?id=ED221530

Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences*, *24*, 11–21.

Dee, T. S. (2004). Teachers, Race, and Student Achievement in a Randomized Experiment. *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.

Deming, D. J., & Noray, K. L. (2018). *Stem careers and technological change.* National Bureau of Economic Research.

Deming, D. J., & Noray, K. L. (2020). Earnings dynamics, changing job skills, and stem careers. *The Quarterly Journal of Economics*, *135*(4), 1965–2005.

Digest of education statistics. pocket digest. (n.d.). National Center for Education Statistics. Retrieved November 21, 2020, from https://nces.ed.gov/ipeds/tablefiles/tableDocs/IPEDS201112Tablesdoc.xlsx

Duprey, M. A., Pratt, D. J., Jewell, D. M., Cominole, M. B., Fritch, L. B., Ritchie, E. A., Rogers, J. E., Wescott, J. D., Wilson, D. H., & International, R. (Eds.). (2018). *High school longitudinal study of 2009 (hsls: 09): Base-year to second follow-up data file documentation. nces 2018*. National Center for Education Statistics.

Dweck, C., & Leggett, E. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256.

Eden, D., & Shani, A. B. (1982). Pygmalion goes to boot camp: Expectancy, leadership, and trainee performance. *Journal of Applied Psychology*, *62*(2), 194.

Ehrenberg, R., & Brewer, D. (1995). Did teachers' verbal ability and race matter in the 1960s? Coleman revisited. *Economics of Education Review*, *14*(1), 1–21.

Fairlie, R. W., Hoffmann, F., & Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*, *104*(8), 2567–91.

Feldman, K. A. (1988). Effective college teaching from the students' and faculty's view: Matched or mismatched priorities? *Research in Higher Education*, *28*(4), 291–329.

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

Fryer Jr., R. G., & Katz, L. F. (2013). Achieving escape velocity: Neighborhood and school interventions to reduce persistent inequality. *American Economic Review*, *103*(3), 232–37.

Gershenson, S., Hart, C., Hyman, J., Lindsay, C., & Papageorge, N. (2018). The Long-Run Impacts of Same-Race Teachers. (10630). http://ftp.iza.org/dp10630.pdf

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

Hanushek, E. A. (1999). Some findings from an independent investigation of the tennessee star experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis*, *21*(2), 143–163.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of human resources*, *39*(2), 326–354.

Hershbein, B. (2013). Worker signals among new college graduates: The role of selectivity and gpa.

High school longitudinal study of 2009 (hsls:09/16), base-year survey. (2009). U.S. Department of Education, National Center for Education Statistics.

Hoffmann, F., & Oreopoulos, P. (2009). A professor like me the influence of instructor gender on college achievement. *Journal of human resources*, *44*(2), 479–494.

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., Rogers, J. E., Jin, Y., & Leinwand, S. (2011). High school longitudinal study of 2009 (hsls: 09): Base-year data file documentation. nces 2011-328. *National Center for Education Statistics*.

Ingles, S., Pratt, D., Herget, D., Burns, L., Dever, J., Ottem, R., Rogers, J., Jin, Y., Leinwand, S., & LoGerfo, L. (Eds.). (2011). *High school longitudinal study of 2009 (hsls:09)*. National Center for Education Statistics.

Jussim, L., & Harber, K. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *9*(2), 131–155.

Jussim, L., & Eccles, J. S. (1992). Teacher expectations: Ii. construction and reflection of student achievement. *Journal of personality and social psychology*, *63*(6), 947.

Kane, T., & Staiger, D. (2008). Estimating teacher impacts on student achievement: an experimental evaluation. (w14067). https://doi.org/http://dx.doi.org/10.3386/w14607

Kane, T., & Staiger, D. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains)*. MET Project. Bill Melinda Gates Foundation.

Kane, T., Taylor, E., Tyler, J., & Wooten, A. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources*, *46*(3), 587–613.

Kimbrough, E. O., McGee, A. D., & Shigeoka, H. (2017). *How do peers impact learning? an experimental investigation of peer-to-peer teaching and ability tracking* (tech. rep.). National Bureau of Economic Research.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, *114*(2), 497–532.

Lavy, V. (2016). What makes an effective teacher? quasi-experimental evidence. *CESifo Economic Studies*, *62*(1), 88–125.

Loury, L., & Garman, D. (1995). College selectivity and earnings. *Journal of labor Economics*, *13*(2), 289–308.

Lusher, L., Campbell, D., & Carrell, S. (2018). Tas like me: Racial interactions between graduate teaching assistants and undergraduates. *Journal of Public Economics*, *159*, 203–224. https://doi.org/https://doi.org/10.1016/j.jpubeco.2018.02.005

Mullainathan, S., & Bertrand, M. (2001). Do people mean what they say? implications for subjective survey data. *American Economic Review*, *91*(2), 67–72.

Oliver, D., Fairlie, R., Millhauser, G., & Roland, R. (2021). Minority student and teaching assistant interactions in stem. *Economics of Education Review*, *83*, 102125.

Papageorge, N., Gershenson, S., & Kang, K. (2019). Teacher expectations matter. *Review of Economics and Statistics*, *46*(1).

Reynolds, D. (2007). Restraining golem and harnessing pygmalion in the classroom: A laboratory study of managerial expectations and task design. *Academy of Management Learning Education*, *6*(4), 475–483.

Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, *40*(3), 411–451.

Rockoff, J., & Speroni, C. (2010). Subjective and objective evaluations of teacher effectiveness. *The American Economic Review*, *100*(2), 261–266.

Rosenthal, L., R.and Jacobson. (1968). Pygmalion in the classroom. *The urban review*, *3*(1), 16–20.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, *125*(1), 175–214.

Sansone, D. (2017). Why does teacher gender matter? *Economics of Education Review*, *61*, 9–18.

Spera, C., & Wentzel, K. R. (2003). Congruence between students' and teachers' goals: Implications for social and academic motivation. *International Journal of Educational Research*, *39*(4), 395–413. https://doi.org/https://doi.org/10.1016/j.ijer.2004.06.006

Stefanou, C., Stolk, J., Prince, M., Chen, J., & Lord, S. (2013). Self-regulation and autonomy in problem- and project-based learning environments. *Active Learn-*

ing in Higher Education, *14*(2). https://doi.org/https://doi.org/10.1177/1469787413481132

Turner, J., Midgley, C., Meyer, D., Gheen, M., Anderman, E., Kang, Y., & Patrick, H. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. *Journal of Educational Psychology*, *94*(1), 88–106.

Urdan, T., & Schoenfelder, E. (2010). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology*, *44*(5), 331–339.

Winters, J. V. (2014). Stem graduates, human capital externalities, and wages in the us. *Regional Science and Urban Economics*, *48*, 190–198.

Tables

Table 1: Key Variables to Analyze Demographic Mismatch - Student Categories

Variable	(1) All	(2) White	(3) Non-White	(4) Black	(5) Male	(6) Female
Outcomes						
9th grade STEM GPA	2.66	2.76	2.51***	2.24***	2.54	2.77***
STEM AP Credits	0.24	0.22	0.27***	0.15***	0.23	0.24
High School STEM GPA	2.47	2.57	2.32***	2.10***	2.34	2.60***
Self-efficacy	0.07	0.05	0.09***	0.21***	0.18	-0.36***
Explanatory Vars.						
Different Sex Teacher	0.49	0.48	0.49	0.47	0.56	0.40***
Different Race Teacher	0.41	0.06	0.93***	0.90***	0.41	0.40
Teacher has STEM degree	0.41	0.27	0.29	0.28***	0.29	0.28
Teachers' Experience	10.5	11	10	10***	10.5	10.4***
Teacher has Master's	0.52	0.52	0.52	0.52	0.52	0.52
Teaching Advanced Courses	0.68	0.67	0.70***	0.71***	0.68	0.67
Number of Certified Teachers	9.01	8.33	10.05***	9.33***	9.12	8.90***
Disruption is a problem	0.42	0.40	0.46***	0.48***	0.43	0.41***
Dropout is a problem	0.43	0.41	0.48***	0.46***	0.44	0.43
Homework (hrs. per day)	1.55	1.53	1.59***	1.62***	1.51	1.61***
Homework (hrs. per day) squared	3.08	2.97	3.24***	3.34***	2.95	3.21***
N	16,730	10,080	6,650	1,360	8,300	8,430

The analytical sample has the structure of an unbalanced panel. Two observations for each course - math and science - are associated with 82% of the students. The sample Teacher's Experience = how many years the teacher taught math/science to 9th to 12th grade students. Teacher has Master's = 1 if the teacher has a Master's or a higher-level degree, 0 if the teacher has a Bachelor's or similar degree. Teaching Advanced Courses = 1 if the most advanced courses are allocated to mostly to the senior teachers. Dropout is a problem = Teacher thinks that students dropping out of class is a problem in this school. Disruption = Teacher thinks that student disruption is a problem in this school. Homework = self-reported hours per day student spends doing homework. Homework squared = previous variable squared.
***p<0.01, **p<0.05, and *p<0.1 in column 3, 4, 6 refer to mean-difference t tests between column 2 and 3, 2 and 4, and 5 and 6 respectively.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 2: Key Variables to Analyze Demographic Mismatch - Teacher Categories

Variable	(1) All	(2) White	(3) Non-White	(4) Black	(5) Male	(6) Female
Outcomes						
9th grade STEM GPA	2.66	2.68	2.50***	2.35***	2.65	2.67
STEM AP Credits	0.24	0.24	0.22**	0.19	0.22	0.25***
High School STEM GPA	2.47	2.49	2.37***	2.32***	2.45	2.49***
Self-efficacy	0.07	0.07	0.10	0.13	0.06	0.08
Explanatory Vars.						
Different Sex Student	0.49	0.48	0.46	0.43***	0.49	0.48**
Different Race Student	0.41	0.37	0.75***	0.74***	0.40	0.42
Teacher has STEM degree	0.28	0.28	0.31***	0.23***	0.30	0.27***
Teachers' Experience	10.5	10.68	8.88***	8.51***	11.2	9.99***
Teacher has Master's	0.52	0.52	0.55***	0.56***	0.52	0.52***
Teaching Advanced Courses	0.68	0.68	0.70***	0.62**	0.66	0.70
Number of Certified Teachers	9.01	8.87	10.25***	9.40***	8.72	9.22***
Disruption is a problem	0.42	0.42	0.49***	0.49***	0.49	0.40***
Dropout is a problem	0.43	0.43	0.51***	0.52***	0.41	0.46
Homework (hrs. per day)	1.55	1.62	1.53***	1.62***	1.60	1.51***
Homework (hrs. per day) squared	3.08	3.07	3.24***	2.80***	3.21	2.95***
N	16,730	15,060	1,670	520	6,970	9,770

There are two observations per student. Teacher's Experience = how many years the teacher taught math/science to 9th to 12th grade students. Teacher has Master's = 1 if the teacher has a Master's or a higher-level degree, 0 if the teacher has a Bachelor's or similar degree. Teaching Advanced Courses = 1 if the most advanced courses are allocated to mostly to the senior teachers. Dropout is a problem = Teacher thinks that students dropping out of class is a problem in this school. Disruption = Teacher thinks that student disruption is a problem in this school. Homework = self-reported hours per day student spends doing homework. Homework squared = previous variable squared.
***p<0.01, **p<0.05, and *p<0.1 in column 3, 4, 6 refer to mean-difference t tests between column 2 and 3, 2 and 4, and 5 and 6 respectively.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 3: Test for Race-based Endogenous Sorting

	(1) NWT	(2) NWT	(3) NWT	(4) NWT
(a) Student non-white	0.00358	0.00431	0.00336	0.00316
	(0.00463)	(0.00466)	(0.00464)	(0.00464)
(b) Teacher-specific MEANTHETA	0.00904	0.00602		
	(0.00541)	(0.00691)		
(c) Interaction between (a) and (b)		0.00517		
		(0.00567)		
(d) Teacher-specific Mean Past Grade (ZG8)			0.0142	0.00684
			(0.00847)	(0.0125)
(e) Interaction between (a) and (d)				0.0128
				(0.0118)
Constant	0.994***	0.993***	0.998***	0.997***
	(0.00363)	(0.00361)	(0.00255)	(0.00264)
N	16730	16730	16730	16730
adj. R^2	0.518	0.518	0.518	0.518

Standard errors in parentheses. Standard Errors are clustered at school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

School × Course fixed effects are included in estimating results in each column.

NWT = Non-white Teacher indicator. I find unique math and science teachers in every school using variables from teacher questionnaire. I create two mean scores specific to each unique teacher. The first mean score variable is Teacher-specific MEANTHETA. Theta Scores are standardized algebra test scores built in the survey. This score has not been seen by the teachers or the schools. I create unique teacher specific theta score. The second mean score variable comes from averaging the past z-scored grade of the highest math/science course the student took in 8th grade over students taught by a unique teacher.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 4: Test for Endogenous Race-based Sorting Affecting Average Teacher-specific Achievement

	(1) MEANTHETA	(2) ZG8	(3) MEANTHETA	(4) ZG8
Non-white Student	-0.00972	0.00655	0.119	-0.00649
	(0.0577)	(0.0371)	(0.0737)	(0.0310)
Non-white Teacher	0.0873	0.0585	0.189	0.00113
	(0.0748)	(0.0495)	(0.6031)	(0.0246)
Non-white Student×Non-white Teacher	0.0637	0.0439	-0.0948	0.0417
	(0.0590)	(0.0386)	(0.0760)	(0.0326)
Constant	0.304***	-0.0757	-0.371***	0.00271
	(0.0750)	(0.0491)	(0.0615)	(0.0220)
N	16730	16730	16730	16730
adj. R^2	0.392	0.008	0.003	0.001

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- School × Course fixed effects are included in estimating results in column 1 and 2. Course fixed effects are included in estimating results in column 3 and 4.
- The two dependent variables MEANTHETA and ZG8 are mean scores specific to each unique teacher. The first mean score variable (MEANTHETA) comes from the theta scores presented in the survey. Theta Scores are standardized algebra test scores built in the survey. This score has not been seen by the teachers or the schools. The second mean score variable (ZG8) comes from averaging the past z-scored grade of the highest math/science course the student took in 8th grade over students taught by a unique teacher.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 5: Test for Selection Bias in *within*-student Panel Model

	(1) F09	(2) F09	(3) F09	(4) F09	(5) F09	(6) F09	(7) F09	(8) F09
ZG8	0.00134	0.00519	-0.00359	-0.0114	-0.00123	0.0174	0.00119	0.00151
	(0.00273)	(0.00406)	(0.00441)	(0.00935)	(0.00707)	(0.0124)	(0.00400)	(0.00398)
Science FE	-0.0372***	-0.0357***	-0.0398***	-0.0500***	-0.0438***	-0.0103	-0.0435***	-0.0310***
	(0.00485)	(0.00554)	(0.00606)	(0.0131)	(0.00846)	(0.0101)	(0.00583)	(0.00560)
Constant	0.924***	0.936***	0.906***	0.894***	0.893***	0.918***	0.921***	0.926***
	(0.00231)	(0.00263)	(0.00284)	(0.00606)	(0.00401)	(0.00571)	(0.00279)	(0.00265)
Sample	All	White	Non-white	Black	Hispanic	Asian	Male	Female
N	27570	15900	11670	2540	3960	2150	13830	13740
adj. R^2	0.015	0.014	0.016	0.020	0.020	0.003	0.019	0.011

Standard errors in parentheses. Standard Errors are clustered at school level. ZG8 = Grade received in the highest math/science course taken in 8th grade. F09 = Binary variable, 1 if student has taken math and science in 9th grade.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 6: Pooled OLS Regression Results - Teachers' Race

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9
Black Teacher	-0.333***			
	(0.0322)			
White Teacher		0.121***		
		(0.0181)		
Hispanic Teacher			-0.221***	
			(0.0260)	
Asian Teacher				0.421***
				(0.0302)
Male Teacher	-0.209***	-0.211***	-0.214***	-0.213***
	(0.0181)	(0.0181)	(0.0181)	(0.0180)
T has a STEM Degree	0.0472***	0.0500***	0.0492***	0.0409**
	(0.0182)	(0.0182)	(0.0180)	(0.0180)
T's Experience of teaching 9-12 grade yrs.	0.00331***	0.00326***	0.00324***	0.00324***
	(0.00101)	(0.00102)	(0.00102)	(0.00102)
T has Master's or higher degree	0.0310**	0.0295*	0.0247	0.0240
	(0.0155)	(0.0157)	(0.0157)	(0.0155)
Science FE	0.0419***	0.0448***	0.0443***	0.0427***
	(0.0107)	(0.0108)	(0.0107)	(0.0105)
Number of Certified Full-time Ts	0.00203**	0.00320***	0.00266**	-0.000769
	(0.00102)	(0.00104)	(0.00104)	(0.00107)
Senior T teaches Advanced Courses	-0.0251	-0.0277*	-0.0254	-0.0174
	(0.0156)	(0.0158)	(0.0158)	(0.0156)
Disruption is a Problem	-0.0541***	-0.0555***	-0.0554***	-0.0552***
	(0.0170)	(0.0171)	(0.0170)	(0.0170)
Dropout is a Problem	-0.0624***	-0.0672***	-0.0687***	-0.0707***
	(0.0167)	(0.0168)	(0.0167)	(0.0165)
Homework hrs. daily	0.0333	0.0337	0.0260	0.00937
	(0.0310)	(0.0309)	(0.0308)	(0.0314)
Homework^2 hrs. daily	-0.00350	-0.00355	-0.00269	-0.000751
	(0.00612)	(0.00609)	(0.00605)	(0.00622)
Constant	0.0708*	-0.0398	0.0790*	0.0646
	(0.0405)	(0.0429)	(0.0408)	(0.0412)
N	16730	16730	16730	16730
adj. R^2	0.025	0.019	0.022	0.029

Standard errors in parentheses.
Standard Errors are clustered at the level of school.
T= Teacher

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 7: Pooled OLS Regression Results - Differences in 9th Grade GPA in Demographic Samples

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9
Asian Teacher	0.0876				0.0934
	(0.0571)				(0.136)
Black Teacher	0.00986	0.202**			
	(0.0473)	(0.0858)			
Hispanic Teacher	0.00780			0.107*	
	(0.0335)			(0.0618)	
White Teacher	-		-0.00383		
	-		(0.0428)		
Male Teacher	-0.0405***	0.000284	-0.0370*	-0.0239	0.0292
	(0.0155)	(0.0511)	(0.0196)	(0.0413)	(0.0511)
T has STEM Degree	0.0439**	0.112*	0.0333	0.0752	0.0254
	(0.0179)	(0.0662)	(0.0228)	(0.0475)	(0.0542)
T's Experience of teaching 9-12 grade yrs.	0.00359***	0.00153	0.00236*	0.00278	0.000461
	(0.00101)	(0.00349)	(0.00127)	(0.00276)	(0.00317)
T has Master's or higher degree	0.0272*	0.00508	0.0251	0.00301	0.110**
	(0.0156)	(0.0566)	(0.0192)	(0.0466)	(0.0541)
Science FE	0.0464***	0.0137	0.0580***	0.0122	0.0549
	(0.0108)	(0.0495)	(0.0146)	(0.0379)	(0.0476)
Number of Certified Full-time Ts	0.00145	0.00768	0.000476	-0.00282	0.00203
	(0.000960)	(0.00551)	(0.00192)	(0.00342)	(0.00384)
Senior T teaches Advanced Courses	-0.0298*	-0.0510	-0.0152	-0.00268	-0.0757
	(0.0158)	(0.0671)	(0.0201)	(0.0441)	(0.0549)
Disruption is a Problem	-0.0640***	-0.0637	-0.0463**	-0.117***	-0.0247
	(0.0169)	(0.0593)	(0.0212)	(0.0446)	(0.0539)
Dropout is a Problem	-0.0634***	0.0984	-0.0827***	0.0586	0.105*
	(0.0164)	(0.0658)	(0.0216)	(0.0497)	(0.0579)
Homework hrs. daily	0.0536*	0.0721	0.0227	0.158*	0.0283
	(0.0310)	(0.100)	(0.0379)	(0.0856)	(0.103)
Homework^2 hrs. daily	-0.00699	-0.00449	0.000981	-0.0395**	-0.0111
	(0.00611)	(0.0186)	(0.00732)	(0.0160)	(0.0208)
Constant	-0.0554	-0.470***	0.0142	-0.240**	0.332**
	(0.0397)	(0.153)	(0.0641)	(0.114)	(0.153)
Sample	All	Black	White	Hispanic	Asian
N	16730	1360	10080	2210	1290
adj. R^2	0.047	0.012	0.002	0.011	0.015

Standard errors in parentheses.
Standard Errors are clustered at the level of school.
In Column 1, "White" is the reference race category. T = Teacher

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 8: Baseline Panel Results of Demographic Mismatch

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9
Other Sex & same Race	-0.0375**	-0.0325*	-0.178*	-0.0203	-0.0503*	-0.0570	-0.399**	-0.699*
	(0.0184)	(0.0186)	(0.102)	(0.0262)	(0.0282)	(0.179)	(0.179)	(0.387)
Same Sex & other race	0.0326	0.0258	-0.0451	0.0194	0.0477	0.0672	-0.163	-0.474*
	(0.0378)	(0.0590)	(0.0704)	(0.0479)	(0.0576)	(0.103)	(0.154)	(0.249)
Other race & Other sex	0.0270	0.0571	-0.0520	0.0172	0.0405	0.0820	-0.263	-0.449*
	(0.0386)	(0.0560)	(0.0711)	(0.0529)	(0.0531)	(0.0987)	(0.163)	(0.254)
Sample	All	White	Non-white	Female	Male	Hispanic	Black	Asian
N	16730	10080	6650	8430	8300	2210	1360	1290
adj. R^2	0.006	0.011	0.006	0.008	0.005	0.006	0.015	0.040

Standard errors in parentheses. Standard Errors are clustered at school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Student ID is the panel variable. Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 9: Baseline Panel Results of Demographic Mismatch - Demographic Stratification

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9	(9) ZGPA9
Other Sex & Same Race	-0.0375**	-0.0134	-0.0525*	-0.218	0.181	-0.609***	0.127	-0.675	-0.831
	(0.0184)	(0.0262)	(0.0284)	(0.192)	(0.359)	(0.215)	(0.303)	(0.508)	(0.511)
Same Sex & Other Race	0.0326	-0.0113	0.0622	0.109	-0.0493	-0.293	0.235	-0.629	-0.290
	(0.0378)	(0.0659)	(0.110)	(0.139)	(0.164)	(0.197)	(0.206)	(0.477)	(0.260)
Other Sex & Other Race	0.0270	0.0613	0.0629	0.157	-0.0391	-0.379*	0.144	-0.548	-0.329
	(0.0386)	(0.0981)	(0.0648)	(0.144)	(0.156)	(0.206)	(0.233)	(0.481)	(0.268)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.006	0.017	0.008	0.029	0.005	0.036	0.032	0.094	0.037

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Student ID is the panel variable. Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 10: Baseline Specification Extended by Including Past Achievement

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9	(9) ZGPA9
Other Sex & Same Race	-0.0365**	-0.0117	-0.0540*	-0.195	0.195	-0.634***	0.189	-0.736	-0.876
	(0.0185)	(0.0261)	(0.0284)	(0.188)	(0.350)	(0.215)	(0.307)	(0.469)	(0.544)
Same Sex & Other Race	0.0319	-0.00215	0.0583	0.124	-0.0401	-0.299	0.260	-0.664	-0.246
	(0.0366)	(0.0638)	(0.108)	(0.137)	(0.159)	(0.191)	(0.206)	(0.443)	(0.232)
Other Sex & Other Race	0.0292	0.0838	0.0687	0.170	-0.0333	-0.394*	0.175	-0.565	-0.271
	(0.0375)	(0.0968)	(0.0634)	(0.143)	(0.152)	(0.205)	(0.234)	(0.447)	(0.240)
ZG8	0.113***	0.145***	0.0879***	0.0966**	0.0799	0.104	0.0671	0.290***	0.207**
	(0.0133)	(0.0254)	(0.0233)	(0.0468)	(0.0500)	(0.0785)	(0.0425)	(0.0536)	(0.102)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.026	0.046	0.020	0.044	0.014	0.052	0.040	0.202	0.076

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7.
ZG8 = Grade the student earned in the highest math or science class in 8th grade; normalized at school level.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 11: Baseline Specification Extended by Including Past Achievement & 8th Grade Course indicators

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9	(9) ZGPA9
Other Sex & Same Race	-0.0282	-0.0169	-0.0397	-0.150	0.228	-0.362	0.0612	-0.881*	-0.952*
	(0.0180)	(0.0263)	(0.0273)	(0.227)	(0.348)	(0.295)	(0.230)	(0.457)	(0.542)
Same Sex & Other Race	0.0362	-0.0453	0.104	0.180	0.00269	-0.214	0.232	-0.660	-0.330
	(0.0368)	(0.0648)	(0.116)	(0.160)	(0.160)	(0.155)	(0.162)	(0.419)	(0.239)
Other Sex & Other Race	0.0342	0.103	0.0561	0.225	0.0506	-0.306*	0.151	-0.610	-0.354
	(0.0374)	(0.0895)	(0.0693)	(0.163)	(0.153)	(0.172)	(0.178)	(0.426)	(0.249)
ZG8	0.113***	0.145***	0.0879***	0.0966**	0.0799	0.104	0.0671	0.290***	0.207**
	(0.0133)	(0.0254)	(0.0233)	(0.0468)	(0.0500)	(0.0785)	(0.0425)	(0.0536)	(0.102)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.029	0.053	0.021	0.085	0.077	0.113	0.053	0.323	0.150

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. Additional controls: (1) ZG8 = Grade the student earned in the highest math or science class in 8th grade; normalized at school level and (2) dummy indicators of highest course in math/science in 8th grade, and (3) interaction terms (1)×(2)

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 12: Baseline Specification Extended by Including 9th grade course indicators

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9	(9) ZGPA9
Other Sex & Same Race	-0.0361*	-0.0120	-0.0542*	-0.209	0.130	-0.538***	0.205	-0.979**	-0.792
	(0.0184)	(0.0260)	(0.0285)	(0.179)	(0.329)	(0.206)	(0.331)	(0.432)	(0.569)
Same Sex & Other Race	0.0306	-0.00755	0.0610	0.160	-0.0600	-0.244	0.259	-0.664*	-0.202
	(0.0365)	(0.0641)	(0.109)	(0.134)	(0.171)	(0.170)	(0.243)	(0.383)	(0.214)
Other Sex & Other Race	0.0273	0.0860	0.0758	0.201	-0.0530	-0.302	0.178	-0.590	-0.222
	(0.0373)	(0.0955)	(0.0650)	(0.141)	(0.161)	(0.184)	(0.276)	(0.387)	(0.220)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.027	0.047	0.023	0.070	0.030	0.090	0.037	0.244	0.089

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. Variables added in addition to the baseline controls: Reference 9th grade Course Indicator = 9th grade Algebra. Other 9th grade course indicators - Geometry, Other Math, Earth & Physical Science, Biology. Sparse course indicators are grouped under Other math or other science.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 13: Baseline Specification Extended by Including Unique Teacher Fixed Effects

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9	(9) ZGPA9
Other Sex & Same Race	-0.0375**	-0.0134	-0.0525*	-0.218	0.181	-0.609***	0.127	-0.675	-0.831
	(0.0184)	(0.0262)	(0.0284)	(0.192)	(0.359)	(0.215)	(0.303)	(0.508)	(0.511)
Same Sex & Other Race	0.0326	-0.0113	0.0622	0.109	-0.0493	-0.293	0.235	-0.629	-0.290
	(0.0378)	(0.0659)	(0.110)	(0.139)	(0.164)	(0.197)	(0.206)	(0.477)	(0.260)
Other Sex & Other Race	0.0270	0.0613	0.0629	0.157	-0.0391	-0.379*	0.144	-0.548	-0.329
	(0.0386)	(0.0981)	(0.0648)	(0.144)	(0.156)	(0.206)	(0.233)	(0.481)	(0.268)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.006	0.017	0.008	0.029	0.005	0.036	0.032	0.094	0.037

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. In addition to the baseline controls, the results in all the above columns are estimated after adding dummy indicators per unique teacher in each school. 67% of the observations correspond with 4 unique teachers in a single school. But the number of unique teachers are up to 19.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 14: Baseline Specification Extended by Including Past Achievement & 8th Grade Course indicators

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9	(9) ZGPA9
Other Sex & Same Race	-0.0282	-0.0169	-0.0397	-0.150	0.228	-0.362	0.0612	-0.881*	-0.952*
	(0.0180)	(0.0263)	(0.0273)	(0.227)	(0.348)	(0.295)	(0.230)	(0.457)	(0.542)
Same Sex & Other Race	0.0362	-0.0453	0.104	0.180	0.00269	-0.214	0.232	-0.660	-0.330
	(0.0368)	(0.0648)	(0.116)	(0.160)	(0.160)	(0.155)	(0.162)	(0.419)	(0.239)
Other Sex & Other Race	0.0342	0.103	0.0561	0.225	0.0506	-0.306*	0.151	-0.610	-0.354
	(0.0374)	(0.0895)	(0.0693)	(0.163)	(0.153)	(0.172)	(0.178)	(0.426)	(0.249)
ZG8	0.113***	0.145***	0.0879***	0.0966**	0.0799	0.104	0.0671	0.290***	0.207**
	(0.0133)	(0.0254)	(0.0233)	(0.0468)	(0.0500)	(0.0785)	(0.0425)	(0.0536)	(0.102)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.029	0.053	0.021	0.085	0.077	0.113	0.053	0.323	0.150

Standard errors in parentheses. Standard Errors are clustered at school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. Additional controls: (1) ZG8 = Grade the student earned in the highest math or science class in 8th grade; normalized at school level and (2) dummy indicators of highest course in math/science in 8th grade, and (3) interaction terms (1)×(2)

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 15: Teachers' Encouragement and Demographic Mismatch

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9
(a) Other Sex & same Race	-0.0296	-0.0249	-0.168*	-0.0105	-0.0446	-0.0786	-0.338*	-0.968**
	(0.0196)	(0.0198)	(0.0997)	(0.0281)	(0.0292)	(0.174)	(0.172)	(0.385)
(b) Same Sex & other race	0.0262	0.00859	-0.0234	0.0118	0.0468	0.0635	-0.135	-0.488*
	(0.0387)	(0.0654)	(0.0701)	(0.0498)	(0.0584)	(0.109)	(0.142)	(0.259)
(c) Other race & Other sex	0.0287	0.0476	-0.0234	0.0146	0.0472	0.0660	-0.216	-0.467*
	(0.0394)	(0.0615)	(0.0714)	(0.0551)	(0.0540)	(0.107)	(0.153)	(0.267)
(d) TEACHER	-0.0544	-0.0463	0.360*	-0.0171	-0.114	0.361*	0.637**	-0.116
	(0.0452)	(0.0475)	(0.188)	(0.0560)	(0.0704)	(0.212)	(0.306)	(0.520)
Interaction (a)×(d)	-0.0238	-0.0223	-0.0464	-0.0245	-0.0159	0.0813	-0.890**	0.815
	(0.0580)	(0.0581)	(0.268)	(0.0726)	(0.0909)	(0.352)	(0.367)	(0.575)
Interaction (b)×(d)	0.100	0.156	-0.332*	0.0828	0.129	-0.373*	-0.459	0.0475
	(0.0677)	(0.163)	(0.189)	(0.0854)	(0.109)	(0.225)	(0.318)	(0.522)
Interaction (c)×(d)	0.0673	0.153	-0.384**	0.0602	0.0834	-0.126	-0.748**	0.0891
	(0.0732)	(0.116)	(0.195)	(0.100)	(0.0969)	(0.237)	(0.352)	(0.533)
Sample	All	White	Non-white	Female	Male	Hispanic	Black	Asian
N	16730	10080	6650	8430	8300	2210	1360	1290
adj. R^2	0.019	0.025	0.023	0.022	0.017	0.026	0.055	0.068

Marginal Effects

TEACHER								
TEACHER at (a) = 1	-0.0782*	-0.0687	0.313	-0.0416	-0.129*	0.443	-0.253	0.699***
	(0.0459)	(0.0463)	(0.199)	(0.0582)	(0.0684)	(0.277)	(0.215)	(0.243)
TEACHER at (b) = 1	0.0461	0.110	0.0280	0.0658	0.0152	-0.0113	0.178	-0.0684
	(0.0523)	(0.161)	(0.0558)	(0.0670)	(0.0914)	(0.116)	(0.129)	(0.116)
TEACHER at (c) = 1	0.0129	0.107	-0.0246	0.0431	-0.0301	0.236*	-0.111	-0.0268
	(0.0600)	(0.117)	(0.0691)	(0.0902)	(0.0765)	(0.140)	(0.178)	(0.131)

Standard errors in parentheses. Standard Errors are clustered at school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. TEACHER = 1 means that the student reported that he/she took 9th grade math or science because a teacher encouraged it.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 16: Demographic Mismatch and AP Credits

	(1) APC	(2) APC	(3) APC	(4) APC	(5) APC	(6) APC	(7) APC	(8) APC	(9) APC
Other Sex & Same Race	-0.00363	-0.00592	0.00237	0.0428	0.161	0.0454	-0.127	-2.050***	0.774
	(0.0151)	(0.0268)	(0.0228)	(0.162)	(0.128)	(0.0657)	(0.141)	(0.266)	(1.102)
Same Sex & Other Race	0.0206	-0.0849	0.0413	0.0804	0.178	-0.0257	-0.210*	-0.404*	0.375
	(0.0371)	(0.0842)	(0.0701)	(0.138)	(0.142)	(0.0389)	(0.117)	(0.235)	(0.375)
Other Sex & Other Race	0.0337	0.0112	0.133	0.120	0.217	-0.0344	-0.233*	-0.528**	0.404
	(0.0346)	(0.0654)	(0.0876)	(0.148)	(0.150)	(0.0559)	(0.133)	(0.256)	(0.384)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.010	0.019	0.011	0.026	0.031	0.080	0.045	0.171	0.044

Standard errors in parentheses. Standard Errors are clustered at school level. APC = Math and Science AP credits. APC varies between math and science for each student because of the panel data structure.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table 17: Demographic Mismatch and High School Math and Science GPA

	(1) HGPA	(2) HGPA	(3) HGPA	(4) HGPA	(5) HGPA	(6) HGPA	(7) HGPA	(8) HGPA	(9) HGPA
Other Sex & Same Race	-0.0253	-0.0560*	0.0212	-0.685***	-0.403*	-0.159	-0.282	-1.288**	0.0243
	(0.0219)	(0.0320)	(0.0297)	(0.241)	(0.223)	(0.236)	(0.232)	(0.558)	(0.425)
Same Sex & Other Race	-0.0429	-0.0532	-0.130	-0.412**	-0.0433	-0.167	-0.263*	-0.806	-0.0791
	(0.0442)	(0.0748)	(0.123)	(0.173)	(0.127)	(0.174)	(0.138)	(0.529)	(0.274)
Other Sex & Other Race	-0.0358	0.0763	-0.00417	-0.377**	-0.0470	-0.140	-0.246	-0.987*	-0.00898
	(0.0451)	(0.120)	(0.0929)	(0.159)	(0.128)	(0.180)	(0.160)	(0.535)	(0.273)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16730	5100	4980	1150	1050	650	7000	650	650
adj. R^2	0.001	0.008	0.003	0.027	0.010	0.031	0.053	0.115	0.038

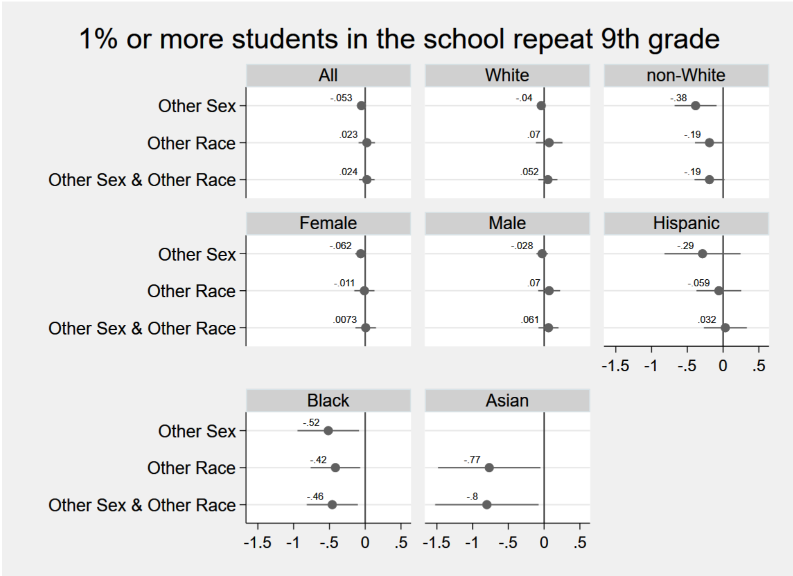
Standard errors in parentheses. Standard Errors are clustered at school level. HGPA = High School GPA. HGPA varies between math and science for each student because of the panel data structure.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

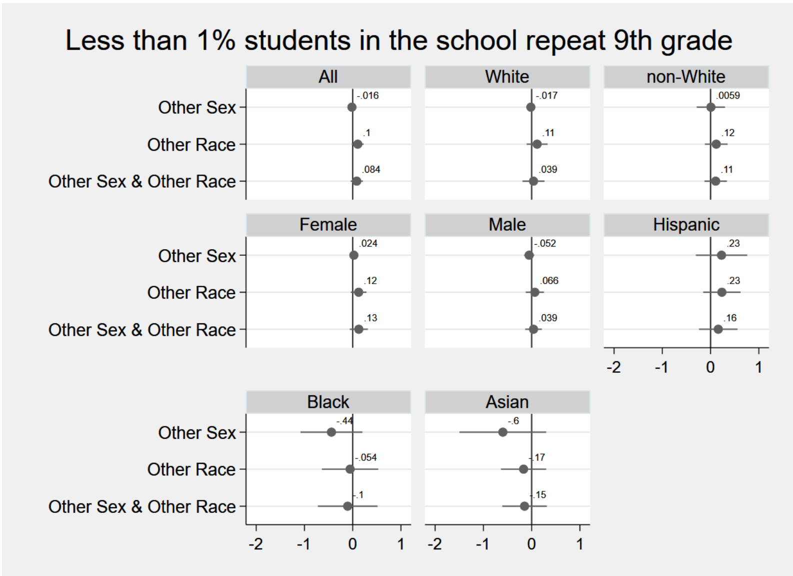
Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Figures

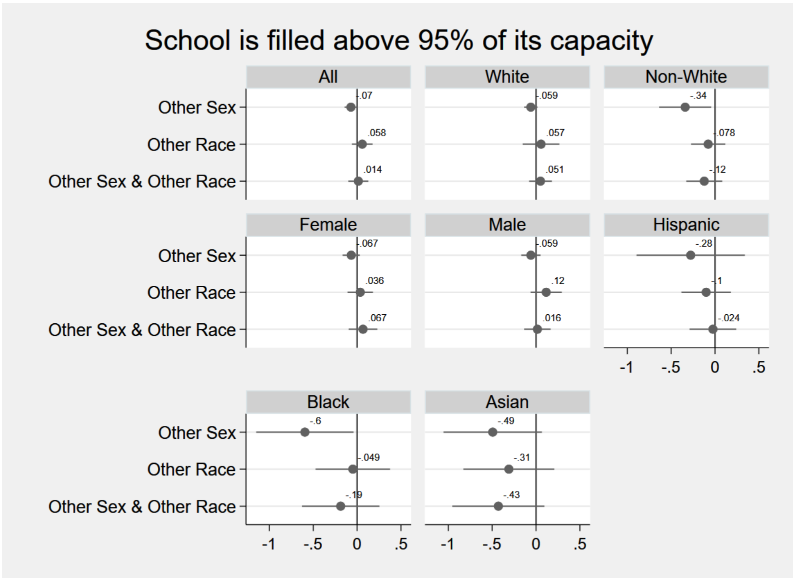


(a)

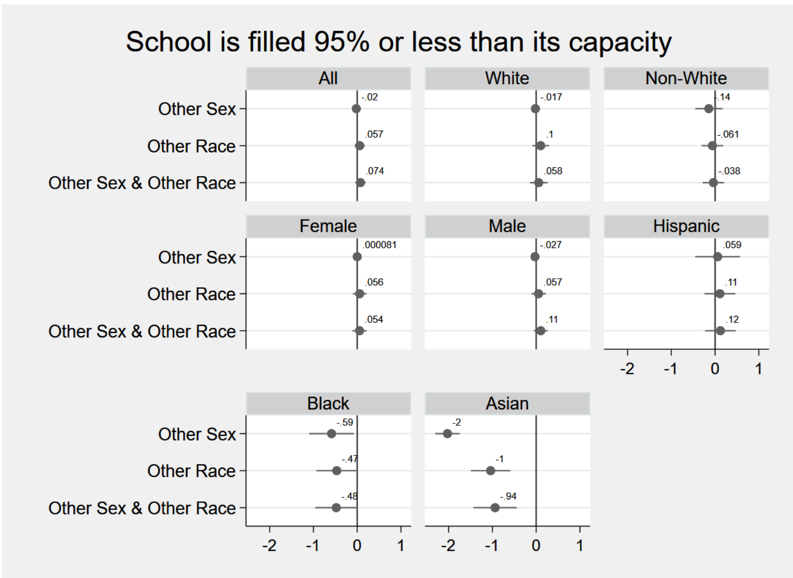


(b)

Figure 1: Percentage of 9th Grade Repeaters and Demographic Mismatch
Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

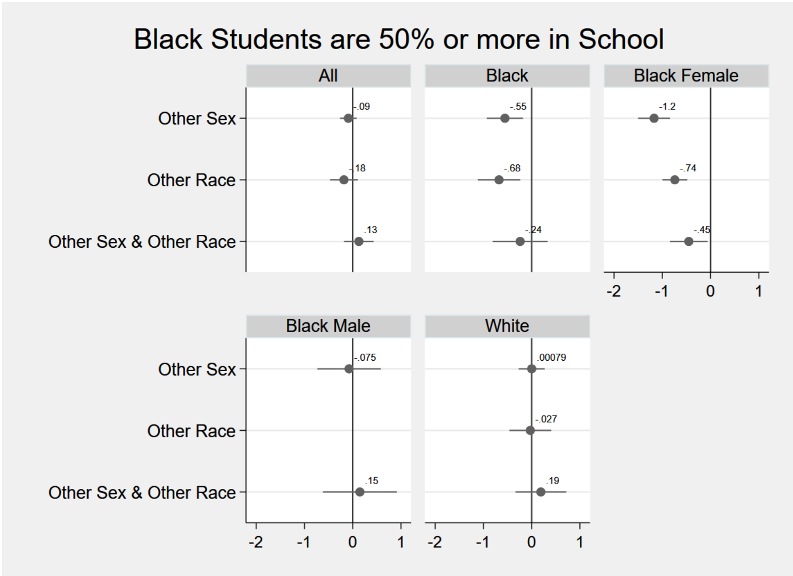


(a)

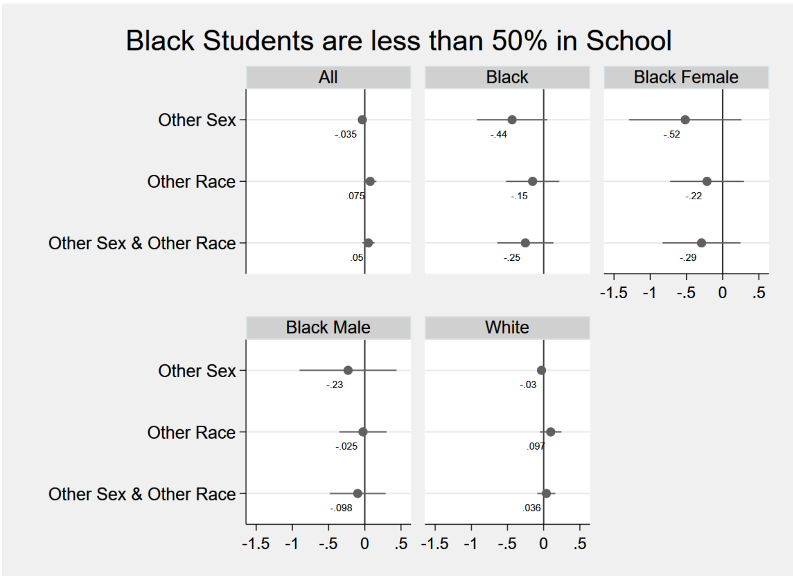


(b)

Figure 2: School Capacity and Demographic Mismatch
Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."



(a)



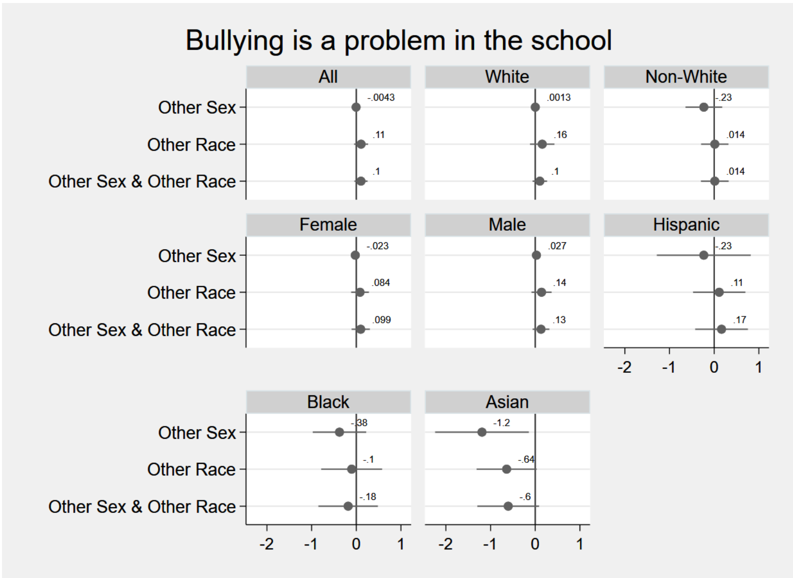
(b)

Figure 3: Percentage of Black Students in a School
Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

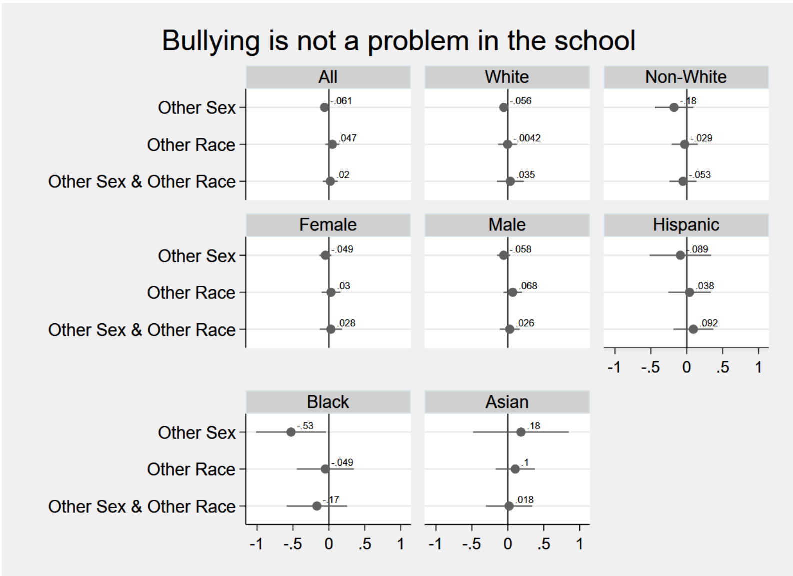
(a)

(b)

Figure 4: Percentage of Students Receiving Free Lunch
Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), “Base-year Survey, 2009.”



(a)



(b)

Figure 5: Whether Bullying is a Problem in the School
Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

A: Supplementary Tables

Table A.1: Additional Variable Means - Past Achievement, Reason for Taking Math/Science, and School Characteristics

Variable	(1) All	(2) White	(3) Non-White	(4) Black	(5) Male	(6) Female
ZGPA8	0.04	0.086	-0.014	-0.015	-0.043	0.011
θ	0.17	0.22	0.10	-0.20	0.168	0.141
	Reason for taking math/science in 9th grade					
Teacher's encouragement	0.11	0.12	0.11	0.079	0.10	0.13
Challenge	0.16	0.16	0.17	0.14	0.16	0.16
Parent's encouragement	0.14	0.15	0.13	0.09	0.13	0.15
Career	0.16	0.15	0.17	0.16	0.14	0.17
Enjoyment	0.16	0.15	0.17	0.16	0.14	0.17
	School characteristics - principal questionnaire					
9th grade repeat	3.92	3.45	4.63	4.89	3.94	3.90
% capacity to which school is filled	87.10	87.08	87.12	85.65	87.33	86.87
% of student body is Black	11.88	9.37	15.68	26.24	11.74	12.02
% of students are enrolled in Free Lunch	30.22	27.79	33.91	33.98	30.13	30.32
Bullying is a problem in the school (=1)	0.49	0.51	0.47	0.45	0.49	0.49
N	16,730	10,080	6,650	1,360	8,300	8,430

ZGPA8 = Z-scored 8th grade highest course GPA, θ = Mathematical theta score.
Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.2: Additional Teacher Survey Controls

Variable	(1) All	(2) White	(3) Non-White	(4) Black	(5) Male	(6) Female
TXFAMILY	0.26	0.26	0.27	0.27	0.27	0.26
TXDISCIPLINE	0.70	0.71	0.69	0.69	0.71	0.69
TXSTUACHIEVE	0.41	0.40	0.43	0.44	0.42	0.40
TXPARENT	0.66	0.66	0.67	0.65	0.66	0.67
TXHOMEFX	0.16	0.15	0.17	0.16	0.16	0.16
TXABLRANGE	0.48	0.48	0.50	0.52	0.49	0.48
TXSESRANGE	0.19	0.18	0.22	0.25	0.19	0.19
Family Income	84,779.16	91,534.52	74,305.67	69,452.69	85,330.63	82,660.69
Parent 1 has College Degree	0.386	0.408	0.355	0.39	0.386	0.386

TXFAMILY = 1 if teacher believes "amount a student can learn is primarily related to family background", 0 if he/she does not; TXDISCIPLINE = 1 if teacher believes "students not disciplined at home are not likely to accept school discipline", 0 if he/she does not; TXSTUACHIEVE = 1 if teacher believes "teachers are limited because home environment influences student achievement", 0 if he/she does not; TXPARENT = 1 if teacher believes "if parents would do more for children teacher could do more for students", 0 if he/she does not; TXHOMEFX = 1 if teacher believes "teacher cannot do much b/c student motivation/performance depends on home", 0 if he/she does not; TXABLRANGE = 1 if teacher believes "teaching is limited by different academic abilities in the same class", 0 if he/she does not; TXSESRANGE = 1 if teacher believes "teaching is limited by students with wide range of socio-economic backgrounds", 0 if he/she does not.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.3: Pooled OLS Regression - Family Characteristics as Controls

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9
Asian T	0.103*				-0.0257
	(0.0577)				(0.130)
Black T	0.0732	0.272***			
	(0.0483)	(0.0962)			
Hispanic T	0.0115			0.0617	
	(0.0373)			(0.0682)	
White T	-		-0.0241		
	-		(0.0449)		
Male	-0.0243	0.0214	-0.0190	-0.0119	-0.0199
	(0.0159)	(0.0532)	(0.0195)	(0.0422)	(0.0542)
T has STEM Degree	0.0200	0.126*	0.00367	0.0415	0.0129
	(0.0183)	(0.0691)	(0.0233)	(0.0497)	(0.0568)
T's Experience of teaching 9-12 grade yrs.	0.00265***	0.00167	0.00156	0.00152	0.00134
	(0.00100)	(0.00364)	(0.00122)	(0.00290)	(0.00337)
T has Master's or higher degree	0.0128	0.00786	0.00505	0.00567	0.102*
	(0.0159)	(0.0602)	(0.0195)	(0.0481)	(0.0555)
Science FE	0.0720***	0.0637	0.0800***	0.0306	0.0563
	(0.0118)	(0.0520)	(0.0150)	(0.0393)	(0.0495)
Number of Certified Full-time Ts	0.00137	0.0118**	-0.000392	-0.00147	0.00118
	(0.00123)	(0.00575)	(0.00211)	(0.00337)	(0.00395)
Senior T teaches Advanced Courses	-0.0154	-0.0583	-0.00740	-0.000549	-0.0306
	(0.0165)	(0.0715)	(0.0203)	(0.0455)	(0.0556)
Disruption is a Problem	-0.0244	-0.0605	0.00321	-0.113**	-0.0380
	(0.0175)	(0.0640)	(0.0214)	(0.0458)	(0.0561)
Dropout is a Problem	-0.124***	0.118	-0.142***	0.0927*	0.0885
	(0.0180)	(0.0720)	(0.0228)	(0.0517)	(0.0602)
Homework hrs. daily	0.0156	0.0481	-0.0229	0.113	0.0349
	(0.0315)	(0.104)	(0.0381)	(0.0890)	(0.109)
Homework^2 hrs. daily	-0.00230	0.000909	0.00691	-0.0334*	-0.0126
	(0.00616)	(0.0189)	(0.00739)	(0.0171)	(0.0222)
Suburb	0.0470***	0.0460**	0.0399	-0.00151	0.0927
	(0.0178)	(0.0192)	(0.0255)	(0.0593)	(0.0667)
Town	0.0887***	0.0849***	0.0909***	0.0313	0.128
	(0.0238)	(0.0260)	(0.0317)	(0.0993)	(0.110)
Rural	0.107***	0.105***	0.0943***	0.0752	0.120
	(0.0198)	(0.0211)	(0.0284)	(0.0611)	(0.0832)
Family Income	1.04e-6***	1.05e-6***	8.59e-7***	1.63e-6***	-2.72e-7
	(1.56e-7)	(1.29e-7)	(1.94e-7)	(4.15e-7)	(4.87e-7)
Parent 1 has College Degree	0.278***	0.278***	0.301***	0.191***	0.202***
	(0.0201)	(0.0167)	(0.0246)	(0.0655)	(0.0614)
Constant	-0.328***	-0.726***	-0.182**	-0.561***	0.00107
	(0.0512)	(0.200)	(0.0812)	(0.146)	(0.177)
Sample	All	Black	White	Hispanic	Asian
N	15490	1230	9410	2010	1190
adj. R^2	0.034	0.014	0.034	0.027	0.013

Standard errors in parentheses. Standard Errors are clustered at school level. In Column 1, "White" is the reference race. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The number of observation is 15490, which is lower than that used for estimating the baseline results. The reason is that the number of missing observations is higher due to the inclusion of family-related variables in the pooled regression model. Family-related variables: family income variable are in the parent survey. The only result of interest in this table is whether the coefficient estimates change by a large margin. The smaller sample does affect the outcome of the sorting test results presented here. These results are using 94% of the analytical sample.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.4: Pooled OLS Regression Results - Teacher Different Race × Other Sex

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9
Teacher Different Race	-0.110***	-0.243**	-0.0214	-0.132	0.0726
	(0.0226)	(0.115)	(0.0583)	(0.0905)	(0.119)
Teacher Different Sex	-0.0375*	-0.190	-0.0406**	-0.0162	0.350
	(0.0202)	(0.126)	(0.0206)	(0.132)	(0.265)
Teacher Different Race × Other Sex	0.00865	0.110	0.0571	0.0553	-0.425
	(0.0307)	(0.135)	(0.0753)	(0.139)	(0.269)
Sample	All	Black	White	Hispanic	Asian
N	16730	1360	10080	2210	1300
adj. R^2	0.004	0.012	0.002	0.010	0.015

Standard errors in parentheses. Standard errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.5: Test for Sorting - Family Income and Parental Educational Attainment

	(1) FAMINCOME	(2) PAR1COL	(3) FAMINCOME	(4) PAR1COL
Non-white Student	17599.9***	0.0680*	22613.3***	0.107**
	(4045.0)	(0.0304)	(4385.6)	(0.0343)
Non-white Teacher	3820.2	0.0624**	10739.5***	0.0843**
	(2762.4)	(0.0210)	(2836.3)	(0.0256)
Ω	406.5	-0.0205	-6804.7	-0.0674
	(4120.2)	(0.0300)	(4584.9)	(0.0351)
Constant	45281.4***	0.121***	63780.2***	0.277***
	(2822.1)	(0.0211)	(2667.7)	(0.0251)
N	15490	15490	15490	15490
adj. R^2	0.216	0.172	0.020	0.005

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Ω is the coefficient to **Non-white Student×Non-white Teacher**. School × Course fixed effects are included in estimating results in column 1 and 2. School fixed effects are included in estimating results in column 3 and 4. The number of observation is 15490, which is lower than that used for estimating the baseline results. The reason is that the main model does not require family-specific variables because student fixed effects are present in the model. Family-related variables: family income and parental income variable are in the parent survey. Parent survey variables have some non-responses which allows estimation with a smaller sample. The only result of interest in this table is whether the coefficient estimate for Non-white Student × Non-white Teacher is statistically significant. The coefficient size is not of great importance. For this reason, a smaller sample does affect the outcome of the sorting test results presented here. These results are using 94% of the analytical sample.
The two dependent variables FAMINCOME and PAR1COL. FAMINCOME is the family income reported by a parent in the survey. PAR1COL is a binary variable with the value 1 when parent 1 has a college degree and 0 when he/she has not.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.6: Demographic Mismatch and Self-efficacy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	EF	EF	EF	EF	EF	EF	EF	EF	EF
Other Sex & Same Race	-0.0567**	-0.0573	-0.0503	-0.0456	-0.225	0.0509	-0.643**	0.257	-1.381**
	(0.0271)	(0.0462)	(0.0399)	(0.185)	(0.295)	(0.361)	(0.323)	(0.317)	(0.578)
Same Sex & Other Race	0.0669	0.0598	0.213	0.162	0.0608	0.142	-0.571***	0.103	-0.895*
	(0.0491)	(0.0926)	(0.132)	(0.155)	(0.234)	(0.319)	(0.148)	(0.235)	(0.509)
Other Sex & Other Race	0.0436	-0.0408	0.0564	0.0515	0.0423	0.127	-0.462***	0.220	-0.992*
	(0.0490)	(0.124)	(0.0907)	(0.154)	(0.229)	(0.313)	(0.171)	(0.259)	(0.506)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	16370	5100	4980	1150	1050	650	700	650	650
adj. R^2	0.009	0.021	0.009	0.024	0.020	0.034	0.101	0.086	0.033

Standard errors in parentheses. Standard Errors are clustered at school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

School × Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.7: Extended Baseline Panel Results - Additional Teacher Survey Instruments as Controls

	(1) ZGPA9	(2) ZGPA9	(3) ZGPA9	(4) ZGPA9	(5) ZGPA9	(6) ZGPA9	(7) ZGPA9	(8) ZGPA9
Other Sex & same Race	-0.0379**	-0.0331*	-0.190**	-0.0194	-0.0488*	-0.0253	-0.437**	-0.758*
	(0.0187)	(0.0190)	(0.0945)	(0.0276)	(0.0280)	(0.155)	(0.182)	(0.409)
Same Sex & other race	0.0231	0.0153	-0.0550	-0.00129	0.0450	0.0829	-0.213	-0.493*
	(0.0391)	(0.0604)	(0.0742)	(0.0495)	(0.0584)	(0.105)	(0.160)	(0.270)
Other race & Other sex	0.0281	0.0702	-0.0524	0.00703	0.0540	0.0896	-0.284*	-0.450
	(0.0396)	(0.0579)	(0.0751)	(0.0533)	(0.0539)	(0.102)	(0.172)	(0.276)
TXFAMILY	0.0109	-0.00414	0.0356	-0.0458*	0.0689**	0.00719	-0.00553	0.219**
	(0.0205)	(0.0251)	(0.0356)	(0.0271)	(0.0292)	(0.0629)	(0.0790)	(0.0947)
TXDISCIPLINE	-0.0101	-0.0128	-0.00513	-0.000779	-0.0190	-0.0245	-0.0335	-0.0521
	(0.0191)	(0.0245)	(0.0302)	(0.0272)	(0.0249)	(0.0526)	(0.0968)	(0.0691)
TXSTUACHIEVE	-0.0262	-0.00922	-0.0557*	-0.00253	-0.0531**	-0.0841	-0.00986	-0.114
	(0.0192)	(0.0233)	(0.0318)	(0.0269)	(0.0264)	(0.0630)	(0.0821)	(0.0765)
TXPARENT	-0.00395	-0.0171	0.0186	-0.0123	0.00165	0.00412	-0.0176	0.0358
	(0.0192)	(0.0234)	(0.0321)	(0.0246)	(0.0265)	(0.0618)	(0.0637)	(0.0706)
TXHOMEFX	-0.00875	-0.0225	0.00847	-0.0255	0.00472	0.0687	0.0700	-0.0598
	(0.0239)	(0.0305)	(0.0395)	(0.0321)	(0.0330)	(0.0684)	(0.0988)	(0.0891)
TXABLRANGE	-0.0227	-0.0378*	-0.000600	-0.0209	-0.0230	0.00118	0.0733	-0.0327
	(0.0173)	(0.0219)	(0.0284)	(0.0237)	(0.0235)	(0.0525)	(0.0622)	(0.0533)
TXSESRANGE	0.0101	0.0308	-0.0220	0.0219	-0.00272	-0.00204	0.0356	-0.0335
	(0.0207)	(0.0251)	(0.0354)	(0.0293)	(0.0288)	(0.0561)	(0.0747)	(0.0996)
Sample	All	White	Non-white	Female	Male	Hispanic	Black	Asian
N	16370	10080	6650	8430	8300	2210	1360	1290
adj. R^2	0.007	0.012	0.005	0.009	0.009	0.010	0.017	0.055

Standard errors in parentheses. Standard Errors are clustered at school level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
School × Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. Additional variables: TXFAMILY = 1 if teacher believes "amount a student can learn is primarily related to family background", 0 if he/she does not; TXDISCIPLINE = 1 if teacher believes "students not disciplined at home are not likely to accept school discipline", 0 if he/she does not; TXSTUACHIEVE = 1 if teacher believes "teachers are limited because home environment influences student achievement", 0 if he/she does not; TXPARENT = 1 if teacher believes "if parents would do more for children teacher could do more for students", 0 if he/she does not; TXHOMEFX = 1 if teacher believes "teacher cannot do much b/c student motivation/performance depends on home", 0 if he/she does not; TXABLRANGE = 1 if teacher believes "teaching is limited by different academic abilities in the same class", 0 if he/she does not; TXSESRANGE = 1 if teacher believes "teaching is limited by students with wide range of socio-economic backgrounds", 0 if he/she does not.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."

Table A.8: Falsification Test: 8th-grade GPA and Future Demographic Mismatch

	(1) ZG8	(2) ZG8	(3) ZG8	(4) ZG8	(5) ZG8	(6) ZG8	(7) ZG8	(8) ZG8	(9) ZG8
Other Sex & same Race	-0.0198	-0.0283	0.0117	0.0739	-0.447	-0.456	-0.732	-0.0162	-0.207
	(0.0228)	(0.0302)	(0.0343)	(0.284)	(0.424)	(0.514)	(0.424)	(0.253)	(0.642)
Other Race & same sex	0.0123	0.0938	-0.0971	-0.246	-0.280	0.0000561	-0.130	-0.0198	-0.249
	(0.0415)	(0.0638)	(0.112)	(0.198)	(0.208)	(0.196)	(0.209)	(0.218)	(0.186)
Other Sex & other Race	-0.0230	-0.161	0.0524	-0.276	-0.249	0.0433	-0.242	-0.140	-0.303
	(0.0460)	(0.9370)	(0.0883)	(0.226)	(0.195)	(0.229)	(0.199)	(0.218)	(0.187)
Science FE	-0.0232*	0.0430	0.0238	-0.119	-0.1000	-0.207*	-0.262	-0.0117	-0.113*
	(0.0128)	(0.0272)	(0.0284)	(0.0735)	(0.0717)	(0.105)	(0.9317)	(0.0811)	(0.0592)
Sample	All	White Female	White Male	Hispanic Female	Hispanic Male	Black Female	Black Male	Asian Female	Asian Male
N	11940	3530	3540	790	770	470	520	520	470
adj. R^2	0.048	0.058	0.049	0.045	0.042	0.048	0.035	0.0332	0.063

Standard errors in parentheses. Standard Errors are clustered at school level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

School × Course fixed effects are included in estimating results in all columns. These results are estimated with the same set of control variables in the regression equation as those in Tables 6 and 7. The samples exclude those students who are in schools that includes an 8th-grade.

Data Source: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09/16), "Base-year Survey, 2009."