Economics of Education Review 52 (2016) 209–224

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Who believes in me? The effect of student–teacher demographic match on teacher expectations

Seth Gershenson[a,b,*], Stephen B. Holt[a], Nicholas W. Papageorge[c]

[a] School of Public Affairs, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016, United States
[b] Institute for the Study of Labor (IZA), Bonn, Germany
[c] Department of Economics, Johns Hopkins University, Wyman Park Building, 3400 N. Charles St., Baltimore, MD 21218, United States

ARTICLE INFO

Article history:
Received 19 August 2015
Revised 25 February 2016
Accepted 3 March 2016
Available online 11 March 2016

JEL classification:
I24
D84
J15
J16

Keywords:
Educational attainment
Expectations
Stigmatization
Mismatch

ABSTRACT

Teachers are an important source of information for traditionally disadvantaged students. However, little is known about how teachers form expectations and whether they are systematically biased. We investigate whether student–teacher demographic mismatch affects high school teachers' expectations for students' educational attainment. Using a student fixed effects strategy that exploits expectations data from two teachers per student, we find that non-black teachers of black students have significantly lower expectations than do black teachers. These effects are larger for black male students and math teachers. Our findings add to a growing literature on the role of limited information in perpetuating educational attainment gaps.

> "You have to ignore it when a child says, 'I don't want to,' because what they're really saying is, 'I don't think I can and I need you to believe in me until I can believe in myself.'"[1]
>
> Shanna Peeples, 2015 CCSSO National Teacher of the Year

1. Introduction

Socio-demographic gaps in educational attainment are well documented (Bailey & Dynarski, 2011; Bound & Turner, 2011). These gaps are especially concerning if they reflect under-investments in human capital among traditionally disadvantaged groups, such as racial minorities or children from low-income families. Sub-optimally low investments in human capital might arise if disadvantaged groups face barriers to educational attainment (e.g., credit constraints).

Limited information, incorrect beliefs, and biased expectations comprise another potentially important, but relatively understudied, source of socio-demographic gaps in educational attainment (Hoxby & Turner, 2013). We examine the formation of public school teachers' expectations of student educational attainment. Teachers likely play an important role in shaping students' beliefs about their academic prospects (Burgess & Greaves, 2013; Dee, 2015), particularly among relatively disadvantaged students who rarely interact with college-educated adults outside of school settings (Jussim & Harber, 2005; Lareau, 2011; Lareau & Weininger, 2008). More concerning, teachers' forecasts can affect students' performance. In a famous

* Corresponding author at: School of Public Affairs, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016, United States.

E-mail address: gershens@american.edu (S. Gershenson).

[1] CCSSO = Council of Chief State School Officers. Quote taken from interview with Envision Education Blog, May 7, 2015. http://www.envisionexperience.com/plan-your-future/blog-articles/congratulations-national-teacher-of-the-year-shanna-peeples.

http://dx.doi.org/10.1016/j.econedurev.2016.03.002

experiment, Rosenthal and Jacobson (1968) manipulated teachers' beliefs of student ability by providing false information regarding students' performance on a nonexistent test and found significantly greater school-year gains among the students who were falsely identified to teachers as "growth spurters". It is troubling, then, that teachers have significantly lower expectations for the educational attainment of socioeconomically disadvantaged and racial minority students (Boser, Wilhelm, & Hanna, 2014). However, whether these "expectation gaps" are evidence of biases in teachers' expectations or simply reflect accurate forecasts (perhaps due to differences in preparation or early childhood investments) is an open question that we address in the current paper.

Specifically, we test for systematic biases in teachers' expectations related to the demographic match between student and teacher using nationally representative survey data in which two teachers reported their expectations for each student's ultimate educational attainment. Differences between two teachers' expectations for the same student may be random in that they reflect idiosyncratic forecasting errors or interactions with a given student. Ex ante, such differences could even be legitimate if they reflect true within-student variation in ability across subjects. For example, if a student excels in math but struggles in reading, the math teacher might correctly forecast a higher level of educational attainment for this student than the student's reading teacher, and vice versa. However, barring a specific type of endogenous sorting of students to teachers – that we later show does not occur – neither of these reasons would explain an association between student–teacher demographic mismatch and within-student differences in teachers' education expectations. Rather, if within-student differences in teachers' expectations are systematically related to the demographic match between student and teacher, this suggests that on average, teachers have systematically biased beliefs about student potential that are at least partly explained by student demographics.

More broadly, large and nationally representative surveys increasingly collect information on subjective beliefs or expectations. For researchers to make causal inferences about how beliefs and expectations affect individuals' decisions and outcomes, they must recognize that beliefs are not only endogenous, but are also potentially biased. Indeed, our key results provide evidence of systematic biases in teachers' expectations. This result highlights the importance of developing and employing credible identification strategies that accurately measure expectations, and biases in expectations, in light of these endogeneity problems when examining the causal relationship between beliefs and economic decision making and outcomes. The identification strategy we propose in this paper borrows heavily from a paper by Dee (2005), which leverages multiple concurrent teacher assessments per student, to implement a student fixed effects strategy. We extend the seminal work of Dee (2005) and subsequent analyses of the impact of student–teacher racial mismatch on teachers' perceptions of student traits and abilities (e.g., Ouazad, 2014) to test for systematic biases in U.S. secondary school teachers' expectations for students' educational attainment.

Expectations are likely correlated with the types of perceptions studied in Dee (2005), e.g., whether a student is frequently disruptive. Still, we argue that evidence of biases in teachers' educational expectations offers important new insights that systematic differences in perceptions cannot. The reason is that the information content is different. Perceptions reflect a teacher's view of a set of a student's characteristics or traits, which may or may not be related to a student's ultimate human capital investments. Expectations questions, in contrast, ask teachers to forecast these investments directly. Therefore, our findings offer direct evidence that demographic mismatch influences how a teacher forms expectations over students' long-run investments. If biased teacher expectations are directly or indirectly communicated to students, they provide precise information about educational investments that perceptions of student traits do not. Precise signals of biased or inaccurate information are worrisome since they could have a relatively large impact on students' own expectations in comparison to information that is less precise.[2] Biased expectations could be incorporated into students' own beliefs, thus influencing their investment decisions. This is especially concerning for disadvantaged students with little prior information on the returns to educational investments. Finally, while a teacher's perceptions reflect their current views of abilities or traits, their expectations are prone to becoming self-fulfilling prophecies if, for example, based on inaccurate forecasts, a teacher shifts scarce resources such as time and effort to another student.

A primary contribution of the current study, then, is to offer guidance to researchers in how to appropriately and fruitfully exploit increasingly available measurements of expectations in large observational data sets. We contribute to the broader literature on the impact of student–teacher racial mismatch along several other dimensions as well, by accounting for more nuanced sources of heterogeneity, such as race-by-gender specific effects. The latter is particularly timely and policy relevant, given recent research documenting the sometimes dramatic sex differences in how disadvantaged children respond to home, school, and neighborhood quality (e.g., Autor, Figlio, Karbownik, Roth, & Wasserman, 2015, 2016; Chetty, Hendren, Lin, Majerovitz, & Scuderi, 2016).

We identify the impact of demographic mismatch on teachers' expectations for students' educational attainment by exploiting a unique feature of the Educational Longitudinal Study of 2002 (ELS): two teachers report their educational expectations for each student. This data structure allows us to condition on unobserved student heterogeneity by making within-student comparisons between the expectations of demographically matched and mismatched teachers. This student fixed effects (FE) identification strategy is motivated by an influential paper by Dee (2005) that exploits a similar feature of the NELS:88

[2] This is based on the idea that, in a standard model with Bayesian updating, noisier signals have a smaller impact on beliefs relative to less noisy, more precise signals since they contain more information (Verrecchia, 1982). Here, the idea is that a teacher's forecast about college-going is a less noisy signal than a teacher's perception that might or might not relate to college-going.

dataset—two teachers appraise the behavior of each student—to identify the effect of demographic mismatch between students and teachers on teachers' perceptions of students' behaviors.[3] Dee finds that when students are assigned to one demographically mismatched teacher and one same-race or same-sex teacher, the demographically mismatched teacher is significantly more likely to perceive the student as being frequently disruptive, frequently inattentive, and less likely to complete homework than is the teacher of a similar demographic background. Similarly, Ehrenberg, Goldhaber, and Brewer (1995) show a robust correlation between student–teacher demographic match and a five-item index that measures teachers' perceptions of students, which includes teachers' binary responses to the question "this student will probably attend college," using the same NELS:88 data.[4] Consistent with Dee (2005) and Ehrenberg et al. (1995), we find that non-black teachers have significantly lower educational expectations for black students than black teachers do. Our results provide insights into the mechanisms through which student–teacher demographic mismatch affects academic achievement and provide novel causal evidence that demographic mismatch affects teachers' expectations for students' long-run educational attainment.[5]

The paper proceeds as follows: Section 2 briefly reviews the relevant theoretical and empirical literatures on biases in teachers' beliefs, stigmatization, and student–teacher demographic mismatch. Sections 3 and 4 describe the data and identification strategy, respectively. Section 5 presents the empirical results and Section 6 concludes.

2. Theoretical background and literature review

Our investigation of the extent to which student–teacher demographic mismatch affects teachers' expectations for students' educational attainment contributes to two distinct literatures. First, a broad, interdisciplinary literature examines biases in beliefs and their impact on decision-making. Mounting evidence suggests that students' beliefs affect their schooling decisions, that their beliefs are often incorrect, and that their beliefs are malleable. For example, Wiswall and Zafar (2015) show that many college students have incorrect beliefs regarding the distribution of average starting salaries across college majors, that students' major choices are a function of these incorrect beliefs, and that biased beliefs can be corrected by an intervention that provides accurate information. Similarly, experimental evidence in social psychology finds that "buffering interventions," which aim to reduce test anxiety attributable to stereotype threat, improve the academic achievement of at least some subsets of the student population (e.g., Dee, 2015; Spitzer & Aronson, 2015).

Biases in beliefs are especially concerning if they lead to under-investments in human capital. For example, a student may forego college if she over-estimates the likelihood of failing to complete her degree. Her decision is sub-optimal in the sense that, given unbiased (accurate) beliefs, she would have matriculated. Morgan, Leenman, Todd, and Weeden (2013) and Dillon and Smith (2013) argue that parents' negatively biased beliefs could lead to under-investment in their children's education, especially in neighborhoods with few college graduates.

Teachers are important inputs in the K-12 education production function who likely shape students' attitudes towards educational attainment (Burgess & Greaves, 2013; Dee, 2015). One channel through which teachers likely influence students' beliefs is via grading (Mechtenberg, 2009). Indeed, robust evidence suggests gender, racial, and ethnic biases in how teachers grade exams in a variety of contexts (Burgess & Greaves, 2013; Cornwell, Mustard, & Van Parys, 2013; Hanna & Linden, 2012; Lavy, 2008). Lavy and Sand (2015) show that grading biases can have long lasting impacts on academic achievement and course taking in high school. Riegle-Crumb and Humphries (2012) study how math teachers stigmatize female students.

Teachers also likely affect students' beliefs by directly imparting their expectations to students. For example, protection models hypothesize that teacher expectations "protect against," or counteract, negative expectations created by neighborhood effects or lack of access to educationally-successful role models (Gregory & Huang, 2013). Indeed, teachers themselves believe that their expectations can affect student outcomes (MetLife, Inc., 2009) and students frequently report favoring teachers who "believe in their ability to succeed" (Curwin, 2012; Golebiewski, 2012). Teachers' expectations strongly predict students' postsecondary educational attainment, though this is not necessarily a causal relationship, as expectations may accurately measure unobservable student ability (Boser et al., 2014; Gregory & Huang, 2013). Still, if teachers' expectations are systematically biased, this likely contributes to the persistence of socio-demographic gaps in educational attainment.

Teachers' expectations might affect student outcomes in at least three ways. First, the perception that teachers have low expectations may exacerbate the harmful effects of *stereotype threat*, whereby low expectations either cause emotional responses that directly harm performance or cause students to *disidentify* with educational environments (Steele, 1997). Second, stigmatized students may modify their expectations, and in turn their behavior, to conform to teachers' negative biases (Ferguson, 2003). In each of the first two cases, teachers' stigmatization of information-poor racial minority students could create a feedback loop that functions like a self-fulfilling prophecy (Burgess & Greaves, 2013; Loury, 2009).[6] Finally, teachers

[3] The NELS:88, or National Education Longitudinal Study of 1998, is a survey conducted by the National Center for Education Statistics that tracked a nationally representative sample of the cohort of U.S. students who were in 8th grade in 1998 over time.

[4] Other binary components of the perceptions index include whether the teacher would recommend the student for honors, thinks the student relates well to others, thinks the student works hard, and interacts with the student outside of class.

[5] Evidence of a causal relationship between student-teacher demographic mismatch and student achievement is accumulating in a variety of school contexts (Antecol, Eren, & Ozbeklik, 2015; Clotfelter, Ladd, & Vigdor, 2007; Dee, 2004, 2007; Egalite, Kisida, & Winters, 2015; Fairlie, Hoffmann, & Oreopoulos, 2014; Lusher, Campbell, & Carrell, 2015).

[6] Stigmatization refers to systematically negatively biased beliefs about a subset of students.

who stigmatize certain types of students may modify how they teach, evaluate, and advise them, again leading to poor educational outcomes for stigmatized students (Ferguson, 2003). All three scenarios potentially perpetuate socio-demographic gaps in educational attainment.

The current study also contributes to the literature on teacher effectiveness. Recent research shows that teachers affect important socioeconomic outcomes including educational attainment, labor market success, and criminal activity (e.g., Chetty, Friedman, & Rockoff, 2014; Jackson, 2012). However, the mechanisms through which high school teachers affect these outcomes are poorly understood. One possible channel is by shaping students' beliefs and expectations about their ability to successfully complete secondary and tertiary education. In that regard, the current study is related to the literature on the relationship between student–teacher demographic mismatch and outcomes such as student test scores, teacher assessments of student behavior and ability, and direct measures of student attendance and suspensions (e.g., Clotfelter, Ladd, & Vigdor, 2007; Dee, 2004, 2005, 2007; Egalite, Kisida, & Winters, 2015; Holt & Gershenson, 2015; McGrady & Reynolds, 2013; Ouazad, 2014). These studies consistently find evidence of arguably causal, modest negative effects of demographic mismatch on academic achievement, teachers' perceptions of student ability, behavior, and non-cognitive skills, and direct measures of student behavior, in both primary and secondary school settings.[7] At the community college level, Fairlie, Hoffmann, and Oreopoulos (2014) find positive effects of being assigned a minority instructor on several measures of minority students' academic success, including course grades, future course selection, and degree completion. However, these studies are typically reduced form in the sense that the mechanisms through which demographic mismatch affects student outcomes are not identified. Teachers' expectations, which may play a particularly important role in shaping the information set used by students and parents to make decisions regarding investments in human capital, are one potential mechanism. The current study investigates this possibility by providing novel evidence of the relationship between student–teacher demographic mismatch and teachers' expectations for students' educational attainment.

3. Data

Data come from the Education Longitudinal Study of 2002 (ELS), which was conducted by the National Center for Education Statistics (NCES). These data are nationally representative of the cohort of U.S. students who were in 10th grade in 2002. Importantly, the ELS elicited subjective expectations of each student's ultimate educational attainment from students' tenth grade math and reading teachers. Having two expectations per student facilitates a within-student identification strategy that we formalize below. The ELS data also contain information on students' demographic and socioeconomic backgrounds, which facilitates analyses of heterogeneous effects of mismatch by student type.

While the data do explicitly link each student to his or her math and reading teachers, the data do not contain unique teacher identifiers, which prevents us from determining with certainty whether two students in a given school were taught by the same teacher.[8] The lack of teacher identifiers is problematic for analyses that require information on which students in a school were assigned to a particular teacher (e.g., the sorting test described below). To facilitate such analyses, we create teacher identifiers using a probabilistic matching process, which is necessarily prone to measurement error. This procedure makes within-school matches based on teachers' race, sex, subject, educational attainment, experience, and college majors and minors and is likely to perform well given the relatively large number of observable teacher characteristics and the fact that the sample is limited to teachers of tenth graders.[9]

Table 1 summarizes the analytic sample of 16,810 student–teacher dyads, each containing exactly two teacher expectations per student, for whom the relevant socio-demographic variables are observed.[10] Column 1 of Table 1 shows that 19% of teachers expected the student to complete no more than a high school diploma while 53% of teachers expected the student to complete a 4-year college degree or more. The categorical ELS expectations variables are more nuanced than those reported in Table 1, but like Dee (2005), we consolidate expectations into "high" and "low" attainment categories to facilitate the estimation of linear and logistic student FE models.[11]

The independent variables of interest measure the degree of demographic mismatch between students and teachers, as characterized by four mutually exclusive categories in column 1: same race and same sex, other race but same sex, same race but other sex, and other race *and* other sex. Overall, about one third of student–teacher pairs are same race and same sex, while another third of the sample is same race but other sex. The remaining third of student–teacher pairs in the analytic sample is similarly evenly split between other race-same sex and other race-other sex pairs. The remainder of column 1 provides information on the observable characteristics of the students and teachers who comprise the analytic sample.

Columns 2 and 3 of Table 1 compare the average characteristics of white and black students, respectively. Black

[7] However, Antecol, Eren, and Ozbeklik (2015) exploit random assignments of Teach for America teachers and find a negative effect of female teachers on female students' math achievement in the most disadvantaged schools.

[8] This is not an issue in analyses of administrative data (e.g., Fairlie et al. 2014).

[9] For example, if two tenth graders in an ELS school had a math teacher who was a white female with a Master's degree, 10 years of teaching experience, and a major in secondary education, we would assume that the two students were taught by the same teacher. To the extent that there are multiple tenth grade teachers in an ELS school who share the same observable profile, this matching process will incorrectly infer that there is only one such teacher.

[10] All sample sizes are rounded to the nearest 10, in accordance with NCES guidelines.

[11] The main results are robust to using alternative definitions of the educational attainment expectations variables (e.g., in correlated random effects ordered-logit models).

Table 1
Analytic sample means by student type (weighted by ELS sampling weights).

Sample T = teacher	All (1)	White students (2)	Black students (3)	Male students (4)	Female students (5)
T expects ≤ HS diploma	0.19	0.15***	0.31	0.23***	0.16
T expects ≥ 4 year degree	0.53	0.58***	0.37	0.49***	0.58
Same race, same sex T	0.34	0.48***	0.10	0.23***	0.45
Other race, same sex T	0.17	0.03***	0.40	0.12***	0.21
Same race, other sex T	0.32	0.45***	0.09	0.43***	0.21
Other race, other sex T	0.17	0.03***	0.40	0.21***	0.12
Reading score	51.01	53.20***	44.64	50.35***	51.67
Math score	51.12	53.35***	43.96	51.79***	50.45
9th grade GPA	2.77	2.90***	2.28	2.65***	2.88
Mom has ≤ HS diploma	0.39	0.35***	0.41	0.37***	0.41
Mom has ≥ 4 year degree	0.26	0.29***	0.20	0.27**	0.25
Low-income HH	0.08	0.05***	0.19	0.07***	0.10
High-income HH	0.14	0.17***	0.06	0.14	0.13
T's experience	14.80	15.15	14.94	14.63	14.96
T has graduate degree	0.51	0.51	0.49	0.50	0.51
T has major in subject taught	0.56	0.58	0.58	0.55**	0.57
N	16,810	10,600	1840	8320	8480

Notes: Student–teacher pairs are the unit of analysis, so there are two observations (teachers) per student. HS = high school. HH = household. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$ refer to mean-difference t tests between columns 2 and 3, and 4 and 5.

students comprise about 11% of the analytic sample. Teachers have systematically lower expectations for black students' educational attainment than for white students. These differences are statistically significant and are consistent with the mean differences in teachers' perceptions of students in the NELS:88 (Ehrenberg et al., 1995). Another notable difference is in the frequency with which white and black students experience demographic mismatch in the classroom, which is due to the majority of teachers being white. White students also have significantly higher test scores, GPAs, and household incomes, which is consistent with evidence of a longstanding racial achievement gap (e.g., Fryer, 2011). Such differences motivate the within-student research design, as the multitude of observed and unobserved differences between white and nonwhite students likely jointly predict teacher expectations and assignment to other-race teachers. Interestingly, however, there are no significant differences between the observable characteristics of teachers assigned to black and white students.

Columns 4 and 5 of Table 1 similarly compare male and female students. On average, teachers have significantly higher expectations for females, which is consistent with the recent reversal of the gender gap in educational attainment (Bailey & Dynarski, 2011; Bound & Turner, 2011) and patterns in teachers' perceptions of students in the NELS:88 (Ehrenberg et al., 1995). There are also significant differences in exposure to other-sex teachers, which is due to the overrepresentation of females in the teaching profession. Otherwise, male and female students come from similar households and are taught by observably similarly teachers.

Table 2 similarly summarizes the analytic sample of 16,810 student–teacher dyads, this time separately by observed teacher characteristics. Columns 1 through 4 of Table 2 present summary statistics separately by teacher demographics. White and female teachers are marginally more optimistic about students' educational outcomes than black and male teachers, respectively. Overall, the white teacher summary statistics strongly resemble those for the full sample, again because the majority of teachers are white. Column 2 shows that black teachers are significantly less likely than white teachers to have same-race students. Black teachers also have significantly lower-performing and lower-SES students than white teachers. These results are consistent with the literature on teacher mobility that finds white teachers are more likely to work in higher-performing, higher-income, suburban schools and that black teachers tend to move to schools with larger black student populations (Hanushek, Kain, & Rivkin, 2004; Jackson, 2009). White teachers are about 8 percentage points more likely to hold a graduate degree than black teachers, and this difference is strongly statistically significant. This highlights the potential importance of controlling for teacher characteristics in the econometric model. There are fewer differences by teachers' sex in the types of students they are assigned, though male teachers are significantly more experienced and more likely to hold a graduate degree than female teachers. The biggest observable difference between male and female teachers is in subject taught. Indeed, columns 5 and 6 show that math teachers are almost twice as likely to be male as are reading teachers. Otherwise, math and reading teachers tend to have similar observable characteristics.

Table 3 presents estimates of descriptive regressions that provide a more nuanced analysis of raw and conditional demographic gaps in teachers' expectations for students' educational attainment. Specifically, Table 3 presents OLS estimates of linear probability models (LPM) in which the dependent variable is a binary indicator equal to one if the teacher expected the student to complete a four-year college degree or more, and zero otherwise.[12] Column

[12] Online Appendix Table A1 presents analogous estimates of linear probability models in which the dependent variable is a binary indica-

Table 2
Analytic sample means by teacher type (weighted by ELS sampling weights).

Sample T = teacher	White teachers (1)	Black teachers (2)	Male teachers (3)	Female teachers (4)	Math teachers (5)	Reading teachers (6)
T expects ≤ HS diploma	0.19**	0.23	0.21***	0.19	0.20	0.19
T expects ≥ 4 year degree	0.54**	0.49	0.52**	0.54	0.53	0.54
Same race, same sex T	0.37***	0.26	0.34	0.35	0.35	0.34
Other race, same sex T	0.14***	0.27	0.18*	0.16	0.17	0.17
Same race, other sex T	0.35***	0.23	0.31***	0.33	0.32*	0.33
Other race, other sex T	0.14***	0.24	0.18***	0.16	0.17	0.16
Reading score	51.51***	46.39	50.94	51.04	51.01	51.01
Math score	51.66***	45.62	51.10	51.14	51.12	51.12
9th grade GPA	2.79***	2.42	2.74**	2.78	2.77	2.77
Mom has ≤ HS diploma	0.38***	0.46	0.38	0.39	0.39	0.39
Mom has ≥ 4 year degree	0.27***	0.21	0.25*	0.26	0.26	0.26
Low-income HH	0.08***	0.18	0.08	0.09	0.08	0.08
High-income HH	0.15***	0.09	0.14	0.14	0.14	0.14
T's experience	15.01**	16.03	15.54***	14.41	15.04**	14.55
T has graduate degree	0.52***	0.44	0.55***	0.48	0.52***	0.49
T has major in subject taught	0.57	0.55	0.57	0.56	0.55***	0.58
T is white	1	0	0.88	0.88	0.87***	0.88
T is black	0	1	0.03***	0.05	0.04*	0.05
T is male	0.35***	0.24	1	0	0.45***	0.25
Math teacher	0.50***	0.46	0.64***	0.42	1	0
Reading teacher	0.50	0.54	0.36***	0.58	0	1
N	14,800	720	5910	10,830	8400	8400

Notes: Student–teacher pairs are the unit of analysis, so there are two observations (teachers) per student. HS = high school. HH = household. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$ refer to mean-difference t tests between columns 1 and 2, 3 and 4, and 5 and 6.

1 reports coefficient estimates for a basic set of demographic characteristics. Relative to the white reference category, teachers are about 20 percentage points less likely to expect black and Hispanic students to complete a college degree but 16 percentage points more likely to expect Asian students to do so. These differences are strongly statistically significant, as is the 9 percentage point gender gap that favors females.

Column 2 of Table 3 reports estimates of a model that also conditions on household SES. Doing so reduces the black–white and Hispanic–white gaps by about 40% and 70%, respectively, which is unsurprising given the lower SES of many black and Hispanic households. The coefficient estimates on the SES indicators in column 2 are of the expected sign and provide evidence of an SES gradient in teacher expectations: teachers have significantly higher expectations for the educational attainment of students from high-income and highly-educated households. The expectation gaps between high- and low-income students and between the children of college-educated and high-school dropout mothers of about 30 percentage points are practically large as well, relative to the unconditional race/ethnicity gaps observed in column 1.

Column 3 of Table 3 further enriches the conditioning set with three measures of academic performance: 9th grade GPA and performance on math and ELA standardized tests. Doing so causes the black and Hispanic point estimates to change signs, indicating that conditional on SES and academic achievement, teachers are significantly more likely to expect traditionally underrepresented minorities to complete a four year college degree than white students. There are at least two possible explanations for this result. First, this might reflect teachers' perceptions of race-based admissions and financial aid policies. Second, this might reflect teachers' beliefs that racial minority students who perform well academically, overcoming perceived challenges in the process, are more motivated than observationally similar white students. It is also notable that the gender gap shrinks by more than 50% after conditioning on academic achievement, though the gap remains statistically significant and in favor of females. It is similarly interesting, and perhaps reassuring, that the SES gradient in teacher expectations significantly flattens after conditioning on students' academic achievement. The coefficient estimates on the academic achievement variables themselves are all of the expected sign and strongly statistically significant.

Finally, column 4 of Table 3 adds school fixed effects to the LPM that control for unobserved school climate and disparities in school and neighborhood resources. Within-school estimates of race/ethnicity indicators are small in magnitude and only the black and Asian coefficient estimates are even marginally statistically significant, though the gender gap remains similar in size and strongly statistically significant. The other point estimates, and the presence of SES gradients in teacher expectations, are robust to the inclusion of school fixed effects. In sum, Table 3 and online Appendix Table A1 suggest that teachers' expectations for students' educational attainment are shaped by students' sex, SES, and academic performance. Importantly, when making within-school comparisons, these factors dominate the effect of race and ethnicity.

Table 4 reports LPM estimates that similarly describe the relationship between observable teacher characteristics and teachers' expectations that the student will complete

tor equal to one if the teacher expected the student to complete a high school diploma or less, and zero otherwise. The qualitative patterns in Appendix Table A1 are similar to those in Table 3, so only the latter is discussed in the main text.

Table 3
Descriptive linear regressions: teacher expects ≥ four-year college degree.

	(1)	(2)	(3)	(4)
Student is white	(Omitted)			
Student is black	−0.22	−0.13	0.10	0.03
	(0.02)***	(0.02)***	(0.02)***	(0.02)*
Student is Asian	0.16	0.17	0.10	0.04
	(0.02)***	(0.02)***	(0.01)***	(0.02)**
Student is Hispanic	−0.17	−0.05	0.08	−0.00
	(0.02)***	(0.02)***	(0.01)***	(0.02)
Student is Native American	−0.28	−0.24	−0.04	−0.02
	(0.06)***	(0.05)***	(0.04)	(0.05)
Student is multiple races	−0.13	−0.08	0.03	−0.00
	(0.03)***	(0.03)***	(0.02)	(0.02)
Student is male	−0.09	−0.11	−0.05	−0.04
	(0.01)***	(0.01)***	(0.01)***	(0.01)***
HH income < $20,000	(Omitted)			
HH income $20,001–35,000		0.06	0.01	0.02
		(0.02)***	(0.01)	(0.01)
HH income $35,001–50,000		0.11	0.04	0.04
		(0.02)***	(0.02)***	(0.01)***
HH income $50,001–75,000		0.17	0.06	0.05
		(0.02)***	(0.02)***	(0.02)***
HH income $75,001–100,000		0.23	0.09	0.06
		(0.02)***	(0.02)***	(0.02)***
HH income ≥ $100,000		0.29	0.12	0.07
		(0.02)***	(0.02)***	(0.02)***
Mom has no HS diploma	(Omitted)			
Mom completed HS		0.07	0.02	0.02
		(0.02)***	(0.02)	(0.02)
Mom has some college		0.13	0.04	0.03
		(0.02)***	(0.01)***	(0.01)*
Mom has ≥ 4 year degree		0.27	0.08	0.05
		(0.02)***	(0.02)***	(0.02)***
Math score			0.01	0.01
			(0.00)***	(0.00)***
Reading score			0.01	0.00
			(0.00)***	(0.00)***
9th grade GPA			0.22	0.25
			(0.01)***	(0.01)***
School fixed effects	No	No	No	Yes
Adjusted R^2	0.04	0.13	0.39	0.45

Notes: $N = 16{,}810$. Parentheses contain standard errors that are robust to clustering at the school level. HH = household. GPA = grade point average. HS = high school. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

a college degree (or more). Column 1 reports estimates of models that do not condition on any student attributes. Multi-race teachers have lower expectations than white teachers, on average, but there are no other significant differences in teacher expectations by race. Male teachers have marginally lower expectations than female teachers. Teachers who have a graduate degree or major in the subject they teach have significantly higher expectations for students' college success, which might be due to such teachers teaching in higher-level courses. There is no evidence of systematic differences between how math and reading teachers evaluate students. The adjusted R^2 in column 1 is only 0.01, however, indicating that teacher characteristics alone explain little of the total variation in teacher expectations.

Column 2 of Table 4 adds controls for students' demographic backgrounds to the regression model, which causes the black and Hispanic teacher indicators to become statistically significant: conditional on student race and sex, black and Hispanic teachers are significantly more likely to expect that students will complete college than white teachers. Controlling for student demographics does not appreciably change the estimated effects of the other teacher characteristics, as teachers with graduate degrees and majors in the subject they teach continue to expect higher levels of educational attainment from their students. Similar results are obtained in columns 3 and 4, which add controls for students' SES and academic performance to the regression models, respectively. Finally, column 5 adds school FE to the LPM, which generally reduces the magnitude of the estimated effects of observable teacher characteristics on teachers' expectations. Nonetheless, the estimated coefficients on the black and Hispanic indicators remain positive and statistically significant, suggesting that there are within-school differences in how teachers of different demographic backgrounds evaluate student potential. Online Appendix Table A2 shows qualitatively similar patterns in teachers' expectations for low educational attainment (i.e., high school diploma or less).

Table 4
Descriptive linear regressions: teacher expects ≥ four-year college degree.

	(1)	(2)	(3)	(4)	(5)
Teacher is white	(omitted)				
Teacher is black	−0.04	0.05	0.07	0.12	0.05
	(0.03)	(0.03)*	(0.03)***	(0.02)***	(0.02)**
Teacher is Hispanic	0.04	0.10	0.11	0.08	0.05
	(0.03)	(0.03)***	(0.03)***	(0.02)***	(0.02)**
Teacher is Asian	0.03	0.03	0.03	0.05	0.03
	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)
Teacher is Native American	−0.15	−0.11	−0.07	−0.04	−0.02
	(0.12)	(0.11)	(0.08)	(0.07)	(0.07)
Teacher is multiple races	−0.10	−0.07	−0.06	−0.01	−0.01
	(0.04)**	(0.04)*	(0.04)*	(0.03)	(0.03)
Teacher is male	−0.02	−0.02	−0.02	−0.01	−0.01
	(0.01)*	(0.01)	(0.01)	(0.01)	(0.01)
Experience	−0.00	−0.01	−0.00	−0.01	−0.00
	(0.00)**	(0.00)***	(0.00)**	(0.00)***	(0.00)***
Experience squared	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)*	(0.00)	(0.00)**	(0.00)**
Graduate degree	0.06	0.05	0.04	0.03	0.00
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)
Math teacher	−0.01	−0.01	−0.01	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Major in subject taught	0.08	0.07	0.06	0.03	0.02
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)**
Adjusted R^2	0.01	0.05	0.14	0.40	0.35
Controls					
Student demographics	No	Yes	Yes	Yes	Yes
Student SES	No	No	Yes	Yes	Yes
Student achievements	No	No	No	Yes	Yes
School fixed effects	No	No	No	No	Yes

Notes: $N = 16{,}810$. Parentheses contain standard errors that are robust to clustering at the school level. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

Table 5 presents estimates of the full specification shown in column 5 of Table 4 separately by student sex and race. Columns 1 and 2 of Table 5 estimate the descriptive regression separately for male and female students, respectively. Interestingly, black teachers have significantly higher expectations for female students than do teachers from other racial and ethnic backgrounds, but no such difference exists in teachers' expectations for male students. Male teachers and math teachers have lower expectations for female students than their female and ELA-teacher counterparts and the differences are marginally significant, but again such differences are not observed among male students. These results are consistent with recent evidence suggesting that some teachers stigmatize female students, particularly in math courses (Lavy & Sand, 2015; Riegle-Crumb & Humphries, 2012).

Columns 3 and 4 of Table 5 report estimates separately for white and black students, respectively. Black teachers have higher expectations for black students than white teachers and this difference is marginally statistically significant. There is no such racial difference in teacher expectations for white students. There are also some large differences in Asian and Native American teachers' expectations for black students, though these cells are quite small and are likely driven by a handful of observations. Math teachers are marginally less likely to expect black students to graduate from college than are ELA teachers, and again no such difference is observed in the subsample of white students. Online Appendix Table A3 presents qualitatively similar results from analogous analyses of "low educational expectations." These results provide suggestive evidence that teachers' expectations are influenced by the interaction between teacher and student demographics, but these descriptive regressions do not disentangle the effect of demographic mismatch from possibly confounding factors such as unobserved student propensity for educational attainment. We present an empirical strategy for doing so below.

4. Identification strategy

The ELS asked each student's tenth grade math (M) and reading (R) teacher how much education they expected the student to complete. Formally, the expectations (E) of student is subject-s teacher are modeled as

$$E_{is} = \alpha_s + \theta_i + \beta \mathbf{x}_{is} + \delta \mathbf{Other}_{is} + \varepsilon_{is}, \forall s \in \{M, R\}, \quad (1)$$

where α is a subject fixed effect (FE) that controls for systematic differences in math and reading teachers' expectations, θ is a student FE that controls for unobserved student characteristics that influence teachers' expectations (e.g., motivation), $\mathbf{x}$ is a vector of observed teacher characteristics that influence their evaluation of students (i.e., experience, graduate degree, major in subject taught), **Other** is a vector of variables that measure the degree of demographic mismatch between teacher and student, and ε represents unobserved idiosyncrasies of the student–teacher

Table 5
Descriptive linear regressions: teacher expects ≥ four-year college degree.

Student subsample	Male (1)	Female (2)	White (3)	Black (4)
Teacher is white	(omitted)			
Teacher is black	0.03	0.09	0.03	0.08
	(0.03)	(0.03)***	(0.03)	(0.05)*
Teacher is Hispanic	0.06	0.04	−0.00	0.02
	(0.03)*	(0.03)	(0.04)	(0.08)
Teacher is Asian	0.03	0.02	0.05	0.22
	(0.04)	(0.03)	(0.04)	(0.09)**
Teacher is Native American	−0.01	−0.01	−0.11	−0.25
	(0.07)	(0.08)	(0.08)	(0.14)*
Teacher is multiple races	−0.05	0.06	−0.01	0.09
	(0.04)	(0.04)	(0.04)	(0.08)
Teacher is male	−0.01	−0.02	−0.01	−0.00
	(0.01)	(0.01)*	(0.01)	(0.03)
Experience	−0.00	−0.01	−0.01	−0.01
	(0.00)	(0.00)***	(0.00)***	(0.00)
Experience squared	0.00	0.00	0.00	0.00
	(0.00)	(0.00)**	(0.00)**	(0.00)
Graduate degree	−0.00	−0.00	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Math teacher	0.00	−0.02	−0.01	−0.04
	(0.01)	(0.01)*	(0.01)	(0.02)*
Major in subject taught	0.02	0.01	0.02	0.01
	(0.01)**	(0.01)	(0.01)**	(0.02)
Adjusted R^2	0.33	0.34	0.35	0.23
N	8,320	8,480	10,600	1,840

Notes: Parentheses contain standard errors that are robust to clustering at the school level. All models condition on the full set of student demographic, SES, and academic performance covariates in addition to school fixed effects. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

dyad that shaped the teacher's expectation for the student.[13]

If data were only available for one student–teacher pair per student, θ would necessarily be included in the error term. The likely endogeneity of θ would bias OLS estimates of $\boldsymbol{\delta}$, the parameter of interest in Eq. (1), as the sorting of teachers and students into classrooms means that unobserved factors such as motivation, innate ability, and barriers to higher education are likely correlated with observed teacher characteristics. However, having two teacher expectations per student allows us to follow Dee (2005) and Fairlie et al. (2014) in estimating Eq. (1) using a student FE strategy that purges such concerns from the model. For simplicity, there are no s subscripts on $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ in Eq. (1), meaning that the baseline model restricts the effects of teacher characteristics and student–teacher demographic mismatch to be homogeneous across subjects. Below, we relax and test this simplifying assumption by interacting the elements of $\mathbf{x}$ and **Other** with the subject FE and testing the joint significance of the interactions.

Following Dee (2005), the baseline model assumes that **Other** contains two elements: binary indicators for "other race" and "other sex." However, to test for multiplicative effects of having both an "other race" *and* "other sex" teacher on teachers' expectations, we also consider a nonparametric specification of **Other** that categorizes the demographic match between teachers and students as one of four possible mutually exclusive scenarios: same race and same sex (omitted reference group), same race and different sex, different race and same sex, and different race *and* different sex. Again following Dee (2005), we examine heterogeneity in the effects of student–teacher demographic mismatch by estimating Eq. (1) separately for subsamples of the student population, as previous research on student–teacher demographic match finds that effects on achievement and teacher perceptions sometimes vary by race and by other observable student characteristics (e.g., Antecol, Eren, & Ozbeklik, 2015; Egalite et al., 2015; Ouazad, 2014).

The baseline model given in Eq. (1) is treated as a linear probability model (LPM). Standard errors are clustered by school, as both teachers and students are nested within schools during students' sophomore year of high school (Angrist & Pischke, 2009). Linear models are preferred despite the binary nature of the dependent variables because they facilitate the inclusion of student FE and the resulting coefficient estimates can be directly interpreted as average partial effects.[14] Nonetheless, we show in online Appendix

[13] It is important to note that θ captures more than student ability and motivation. It captures any fixed characteristics that might affect teachers' expectations, including a teacher's perception that a student may struggle to complete their schooling (e.g., perceived barriers to higher education due to family circumstances).

[14] Student fixed effects (FE) cannot be included in pooled probit or logit models due to the incidental parameters bias that arises when there are only two observations per student. Meanwhile, it is impossible to estimate proper average partial effects in the conditional (FE) logit model because the distribution of fixed effects is unknown (Wooldridge, 2010).

Table A4 that the baseline results are robust to specifying a nonlinear conditional (FE) logit model that acknowledges the binary nature of the dependent variables, which takes the right hand side of Eq. (1) as its linear index. Another sensitivity analysis is presented in online Appendix Table A5, which shows that the preferred LPM estimates are robust to weighting by NCES-provided sampling weights, as suggested by Solon, Haider, and Wooldridge (2015).

The remaining threat to the validity of the baseline student-FE estimates is endogenous sorting that systematically varies by subject *and* student background. Intuitively, this concern is analogous to those about time-varying unobserved heterogeneity in panel data settings in which individuals are observed repeatedly over time and time-invariant individual FE fail to purge time-varying sources of bias from the model. In the current context, the student FE only adequately controls for sorting into classrooms based on student unobservables if the sorting mechanism is the same for both math and reading classrooms. For example, baseline estimates of $\boldsymbol{\delta}$ in Eq. (1) would overstate the effect of student–teacher demographic mismatch if low math ability nonwhite students are systematically assigned to white math teachers. While we can neither account nor test for such sorting on unobservables, we conduct two exercises using observable data that mitigate concerns that the main results are biased by sorting on unobservables. First, we augment Eq. (1) to control for students' course-*s* grade in 9th grade. Intuitively, 9th grade subject-specific grades proxy for a combination of subject-specific aptitude and effort level. Online Appendix Table A6 shows that the baseline estimates are remarkably robust to this addition, and thus that the results are not driven by demographically mismatched teachers being assigned to students with systematically lower subject-specific prowess. Second, we follow Fairlie et al. (2014) in testing for analogous types of differential sorting on observables, who argue that if there is no systematic sorting on observable student characteristics ($\mathbf{z}$) and the elements of $\mathbf{z}$ are highly correlated with the ε in Eq. (1), then differential sorting on unobservables of the sort described above is unlikely to seriously threaten the validity of the preferred student-FE estimates of Eq. (1).

Implementing a Fairlie et al. (2014) style test for differential sorting by observables requires using teacher identifiers that we created via probabilistic matching to compute $\bar{z}^r_{jk}$, the mean value of characteristic z among teacher j's type-r students in school k, where r could denote race or sex. In the simplest form of the test r is a binary indicator equal to one for nonwhite students and zero otherwise. We then use two observations per teacher to estimate linear regressions of the form

$$\bar{z}^r_{jk} = \omega_k + \lambda 1\{r = 1\} + \pi 1\{r = 1\} \times Nonwhite_{jk} + u^r_{jk}, \quad (2)$$

where ω is either a school or school-by-subject FE, $1\{.\}$ is the indicator function, *Nonwhite* is a binary indicator equal to one if teacher j is nonwhite and zero otherwise, and u is an idiosyncratic error term. The coefficient of interest in Eq. (2) is π, which is essentially a difference-in-differences estimate of how the mean difference between white and nonwhite student characteristics varies between white and nonwhite teachers in the same school (or school-subject pair). If the OLS estimate of π is statistically indistinguishable from zero, there is no evidence of differential sorting on observables and thus differential sorting on unobservables in a way that would bias the preferred student-FE estimates of Eq. (1) is unlikely.

5. Results

5.1. Sorting test results

Table 6 reports estimates of two versions of the sorting test proposed by Fairlie et al. (2014) and described in Eq. (2). Panel A of Table 6 reports estimates from models that condition on school FE and panel B reports estimates from models that condition on school-by-subject FE. The two specifications produce nearly identical results, which alleviates concerns that the main results are biased by differential sorting by subject and student race into classrooms. Specifically, the interaction terms that represent differential sorting by race on observables are statistically indistinguishable from zero for each of the five student characteristics considered: 9th grade GPA, mother has a high school diploma or less, mother has a college degree or more, student comes from a low-income household, and student comes from a high-income household. Moreover, the interaction term point estimates and standard errors are small in magnitude, again suggesting that there is no differential sorting on observables by student race. Together with the results presented in online Appendix Table A6, the sorting test results discussed above suggest that differential sorting on unobservables is unlikely to pose a serious threat to identification, as previous performance, household income, and maternal education are likely correlated with the idiosyncratic error term in Eq. (1).

5.2. Main results

Table 7 reports estimates of the baseline LPM shown in Eq. (1). The first row reports estimates for the full analytic sample and each subsequent row reports estimates for a specific subsample of interest. Columns 1 and 2 report the estimated effects of *Other Race* and *Other Sex* student–teacher pairings, respectively, on the likelihood that teachers expect students to complete a high school diploma or less. In the full sample, the other-race effect is positive,

Another advantage of the LPM is that it can be augmented to include two-way student and teacher FE (e.g., Mittag, 2012). In this specification, teacher FE replace the teacher characteristics contained in $\mathbf{x}$. While exploiting within-teacher variation in this way is appealing, our ability to do so is limited by two practical issues. First, as discussed in Section 3, we must use imputed, potentially incorrect, teacher identifiers. Second, two-way FE estimators can only be implemented for the subsample of teachers who taught multiple students and for whom there is variation in E and **Other**. As a result, the two-way FE analysis is underpowered and yields imprecise estimates (e.g., this restriction cuts the black subsample in half). Estimates of the baseline student FE model on the restricted "two-way FE sample" are similarly imprecise, thus we do not report or attempt to interpret the two-way FE estimates.

Table 6
Sorting test estimates.

	9th grade GPA (1)	Mom has HS or less (2)	Mom has college + (3)	Low income (4)	High income (5)
A. School FE estimates					
Nonwhite Teacher	0.03	−0.02	0.04	−0.03	0.00
	(0.06)	(0.04)	(0.03)	(0.02)*	(0.03)
Nonwhite Student Indicator	−0.15	0.09	−0.06	0.06	−0.07
	(0.03)***	(0.02)***	(0.02)***	(0.01)***	(0.01)***
Interaction term (π)	**−0.09**	**0.01**	**−0.02**	**0.04**	**0.01**
	(0.07)	**(0.05)**	**(0.05)**	**(0.03)**	**(0.04)**
B. School-by-subject FE estimates					
Nonwhite Teacher	0.06	−0.05	0.07	−0.04	0.02
	(0.08)	(0.04)	(0.04)	(0.02)*	(0.03)
Nonwhite Student Indicator	−0.15	0.09	−0.05	0.06	−0.07
	(0.03)***	(0.02)***	(0.02)***	(0.01)***	(0.01)***
Interaction Term (π)	**−0.09**	**0.01**	**−0.02**	**0.04**	**0.01**
	(0.07)	**(0.05)**	**(0.05)**	**(0.03)**	**(0.04)**

Notes: $N = 3030$. Each regression contains two observations per teacher: the white and nonwhite student mean characteristics. FE = fixed effects. GPA = grade point average. HS = high school. Bold interaction terms are the interaction between the nonwhite teacher and nonwhite student mean indicators, which constitute the sorting test described by Eq. (2) in the text. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$.

Table 7
Baseline linear probability model (LPM) estimates of teachers' expectations.

Outcome	≤High School Diploma		≥4 year degree	
Independent variable / Sample	Other race (1)	Other sex (2)	Other race (3)	Other sex (4)
Full sample	0.03	0.01	0.01	−0.01
[N = 16,810]	(0.02)*	(0.01)	(0.02)	(0.01)
White sample	−0.00	0.00	0.03	−0.01
[N = 10,600]	(0.02)	(0.01)	(0.02)	(0.01)
Black sample	0.12	0.00	−0.09	−0.05
[N = 1840]	(0.04)***	(0.03)	(0.05)**	(0.03)*
Hispanic sample	0.03	0.03	0.00	0.01
[N = 2110]	(0.04)	(0.02)	(0.06)	(0.02)
Male sample	0.03	−0.00	−0.00	0.00
[N = 8320]	(0.02)	(0.01)	(0.03)	(0.01)
Female sample	0.03	0.02	0.01	−0.01
[N = 8480]	(0.02)	(0.01)	(0.02)	(0.01)
White male sample	−0.01	0.00	0.00	0.01
[N = 5270]	(0.03)	(0.01)	(0.03)	(0.02)
White female sample	0.00	0.00	0.05	−0.00
[N = 5330]	(0.02)	(0.01)	(0.03)*	(0.02)
Black male sample	0.14	−0.05	−0.11	−0.07
[N = 860]	(0.05)**	(0.04)	(0.06)*	(0.04)*
Black female sample	0.09	0.04	−0.07	−0.03
[N = 980]	(0.05)*	(0.03)	(0.04)*	(0.04)
Low-income sample	0.09	0.01	−0.07	−0.01
[N = 1370]	(0.04)**	(0.03)	(0.05)	(0.03)
High-income sample	0.03	−0.01	0.06	0.00
[N = 2770]	(0.03)	(0.01)	(0.04)	(0.02)
Northeast sample	−0.07	0.01	−0.01	0.02
[N = 3010]	(0.06)	(0.01)	(0.06)	(0.02)
Midwest sample	−0.02	0.02	0.02	−0.01
[N = 4530]	(0.03)	(0.01)*	(0.04)	(0.01)
South sample	0.06	−0.00	0.01	−0.01
[N = 6420]	(0.02)***	(0.01)	(0.02)	(0.01)
West sample	0.02	0.00	−0.03	−0.03
[N = 2850]	(0.02)	(0.01)	(0.04)	(0.02)*

Notes: Each row of columns 1 and 2 reports coefficient estimates from the same regression, and similarly for columns 3 and 4. Parentheses contain standard errors that are robust to clustering at the school level. All models condition on student fixed effects and control for teacher characteristics. There are two observations per student, one each from the student's math and reading teacher.

*** $p < 0.01$,
** $p < 0.05$, and
* $p < 0.1$.

small in magnitude, and only marginally statistically significant. This suggests that on average, teachers are more likely to expect low levels of educational attainment for students of different racial backgrounds than they are for students of the same race.

However, restricting the effect of racial mismatch to be constant across all students might mask important heterogeneities by student race, sex, and SES. Indeed, the subsequent five rows of Table 7 show that the overall positive effect of racial mismatch on the probability that the teacher has low expectations for educational attainment was almost entirely driven by non-black (mostly white) teachers' expectations for black students relative to the expectations of black teachers. Specifically, non-black teachers are 12 percentage points more likely to expect black students to complete a high school diploma or less than are black teachers and this difference is statistically significant. It is also arguably practically significant, as it represents an almost 40% increase relative to the mean expectation for black students of 0.31.

Further stratification of the race/ethnicity subsamples reveals that the effect of student–teacher racial mismatch on teachers' expectations that black students complete a high school diploma or less is five percentage points larger for black males than for black females, though these estimates are less precisely estimated, which is likely due to the substantial reductions in sample size. Other-race teachers are also relatively more likely to have lower educational expectations for students from low-income households and students in the South.

Column 2 of Table 7 provides no evidence of an effect of sex mismatch on teacher expectations for low educational attainment, either overall or by student subgroup.

Columns 3 and 4 of Table 7 similarly report estimates of the baseline LPM for the probability that teachers expect the student to complete a 4-year college degree or more. Neither the other-race nor other-sex indicator is significant when the model is estimated using the full analytic sample, though as discussed above there might be significant differences in the effect of demographic mismatch by students' demographic and socioeconomic backgrounds. Sure enough, and consistent with the results for low educational expectations presented in column 1 of Table 7, non-black teachers are significantly less likely to expect black students to complete a 4-year college degree than are black teachers. Again, the effect of racial mismatch on teachers' expectations for college completion is larger in magnitude for black males than black females.

Interestingly, and unlike in the results for low expectations reported in column 2, a marginally significant effect of gender mismatch on teachers' expectations for black students' college success is observed in column 4 of Table 7. This appears to be mostly driven by female teachers' expectations of black male students. These results suggest nonlinearities in the effects of other-race and other-sex student–teacher assignments on teachers' expectations, which we further investigate in Table 8.

Specifically, Table 8 reports estimates of a richer version of the preferred LPM in which **Other** is specified as a set of four mutually exclusive categorical indicators of the nature of the demographic match between students and teachers. Same race and same sex is the omitted reference group to which reported point estimates can be compared. In the full analytic sample, the point estimate in column 3 of the first row of Table 8 shows that the overall other-race effect observed in column 1 of Table 7 was driven by instances of racial mismatch in which there was also sex mismatch. In the black subsample, the "other race and other sex" indicator in column 3 of Table 8 is positive and relatively large, but imprecisely estimated. However, the "other sex" indicator in column 2 is negative, twice as large in magnitude, and statistically significant. This indicates that black teachers assigned to black students of the opposite sex are significantly less likely to have low expectations than a black teacher of the same sex as the student. This result is likely driven by black female teachers' expectations for black male students, as black male teachers are relatively rare in the analytic sample. Indeed, in the black-male subsample black female teachers are 20 percentage points less likely than white teachers of either sex, and almost 30 percentage points less likely than black male teachers, to expect a high school diploma or less. In other words, black female teachers are significantly more optimistic about black males' ability to complete high school than teachers from any other demographic group. As for black female students, there is a marginally statistically significant "other race and other sex" effect on teacher expectations, which suggests white male teachers are about 10 to 20 percentage points more likely to have low expectations for black female students than teachers from other demographic backgrounds.

Columns 4–6 of Table 8 similarly analyze the effect of student–teacher demographic mismatch on the probability that teachers have high expectations for students' educational outcomes. Several of the patterns reverse. Notably, among teachers of black students, other-race teachers are significantly less likely to expect a four year college degree, regardless of the sex match between student and teacher. This is in stark contrast to the results for low expectations, and highlights the nuanced relationship between student–teacher demographic mismatch and teachers' expectations for educational success.

Finally, in Table 9 we relax the assumption that the coefficients on demographic mismatch (δ) and observed teacher characteristics ($\boldsymbol{\beta}$) do not vary by subject by augmenting the baseline LPM, which only included a math-teacher FE, to include interactions between the math-teacher FE and each of the model's covariates. These interaction terms are only jointly statistically significant when the augmented model is estimated on the subsample of black students, as shown by the joint *F*-test *p* values reported in column 7 of Table 9, suggesting that the baseline model is a reasonable specification. Still, the role of student–teacher demographic mismatch in shaping teachers' expectations for student attainment might vary between math and reading classrooms.

Panel A of Table 9 investigates whether this is so for the formation of low-attainment (high school or less) teacher expectations. The math-other race interaction term in the first row of panel A is statistically significant and suggests that the small, positive effect of racial mismatch on low teacher expectations in the full sample observed in

Table 8
Four-category linear probability model (LPM) estimates of teachers' expectations.

Outcome	≤High School Diploma			≥4 year degree		
Independent variable Sample	Other race (1)	Other sex (2)	Other race and other sex (3)	Other race (4)	Other sex (5)	Other race and other sex (6)
Full sample	0.01	0.00	0.04[c]	−0.01	−0.01	−0.01
[N = 16,810]	(0.02)	(0.01)	(0.02)*	(0.02)	(0.01)	(0.02)
White sample	−0.02	0.01	−0.00	0.01	−0.01	0.04
[N = 10,600]	(0.04)	(0.01)	(0.04)	(0.03)	(0.01)	(0.04)
Black sample	0.03[a]	−0.15[a]	0.08	−0.12	−0.06	−0.17
[N = 1840]	(0.06)	(0.08)**	(0.07)	(0.06)**	(0.08)	(0.06)***
Hispanic sample	0.04	0.09	0.07	0.00	−0.11	0.01
[N = 2110]	(0.05)	(0.08)	(0.05)	(0.10)	(0.12)	(0.08)
Male sample	0.04	0.01	0.02	−0.00	−0.00	−0.01
[N = 8320]	(0.03)	(0.02)	(0.03)	(0.02)	(0.01)	(0.04)
Female sample	−0.00	0.00[b]	0.06[b]	0.01	−0.02	−0.01
[N = 8480]	(0.03)	(0.01)	(0.03)**	(0.03)	(0.02)	(0.03)
White male sample	0.00	0.01	−0.02	−0.03	0.01	0.02
[N = 5270]	(0.08)	(0.02)	(0.04)	(0.05)	(0.02)	(0.05)
White female sample	−0.03	0.00	0.02	0.04	−0.01	0.09
[N = 5330]	(0.03)	(0.01)	(0.07)	(0.03)	(0.02)	(0.06)
Black male sample	−0.07[a]	−0.27[b]	−0.07	−0.18	−0.12	−0.25
[N = 860]	(0.11)	(0.11)**	(0.11)	(0.14)	(0.15)	(0.15)
Black female sample	0.06	−0.11[b]	0.16[b]	−0.09	0.02	−0.10
[N = 980]	(0.07)	(0.10)	(0.08)*	(0.05)*	(0.14)	(0.06)*
Low-income sample	0.06	−0.01[c]	0.10	−0.05	0.00	−0.02
[N = 1370]	(0.06)	(0.05)	(0.07)	(0.07)	(0.05)	(0.07)
High-income sample	0.05[c]	−0.01	0.01	−0.06	−0.02[c]	0.08[a]
[N = 2770]	(0.04)	(0.01)	(0.03)	(0.05)	(0.03)	(0.06)
Northeast sample	−0.09	0.02	−0.08	0.09	0.02	0.10
[N = 3010]	(0.08)	(0.02)	(0.07)	(0.06)	(0.02)	(0.07)
Midwest sample	−0.03	0.02	0.03	0.01	−0.02[c]	0.08[c]
[N = 4530]	(0.05)	(0.02)	(0.04)	(0.05)	(0.02)	(0.06)
South sample	0.05[b]	−0.03[a]	0.08	0.00	−0.00[c]	−0.05
[N = 6420]	(0.03)*	(0.02)*	(0.03)**	(0.03)	(0.02)	(0.03)
West sample	−0.04	0.01	−0.04	−0.06	−0.07	−0.06
[N = 2850]	(0.04)	(0.02)	(0.04)	(0.06)	(0.04)*	(0.05)

Notes: Each row of columns 1, 2, and 3 reports coefficient estimates from the same regression, and similarly for columns 4, 5, and 6. The omitted mismatch category in each regression is "same race and same sex." Parentheses contain standard errors that are robust to clustering at the school level. All models condition on student fixed effects and control for teacher characteristics. There are two observations per student from the students' math and reading teacher. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.1$; a, b and c similarly indicate the significance of differences between mismatch categories (column 1(4) vs. 2(5), column 2(5) vs. 3(6), and column 3(6) vs. 1(4)).

Table 7 was driven by lower expectations among racially mismatched math teachers as opposed to reading teachers. However, while the analogous interaction term in the black subsample is positive and of the same magnitude, it is statistically indistinguishable from zero. In other words, for black students, the other-race effect on teacher expectations was approximately the same for both math and reading teachers. Finally, the last row of panel A of Table 9 shows that demographically mismatched math teachers are significantly more likely to expect low attainment for female students, while there is no effect of demographic mismatch on reading teachers' expectations for female student attainment. This is consistent with evidence that female students are stigmatized in math classrooms (e.g., Lavy & Sand, 2015).

Panel B of Table 9 does the same for high-attainment expectations (college degree or more). Like in the main results, the only significant effects of demographic mismatch on high expectations are observed in the subsample of black students. Interestingly, however, the other-race coefficient is negative but not statistically significant at traditional confidence levels. In the context of the augmented interaction model, this means that there is no significant effect of racial mismatch on reading teachers' expectations for student attainment. The other race-math interaction term is also negative and statistically insignificant, but combined with the other-race effect, the total effect of racial mismatch on math teachers' expectations of −0.15 is relatively large in magnitude and strongly statistically significant.[15] This suggests that the general finding that racial mismatch between students and teachers lowered teachers' expectations that students would earn a four-year college degree was largely driven by math teachers' expectations.

6. Discussion

Using unique, nationally representative survey data that contain two teachers' expectations for each student's educational attainment, we estimate student fixed-effects models that identify the effect of student–teacher demographic mismatch on teachers' expectations. The estimates are arguably causal, as the identifying variation comes

[15] Standard errors of the net effects were computed via the delta method.

Table 9
Heterogeneous linear probability model (LPM) estimates of teachers' expectations by subject.

Independent variable	Other race (1)	Math × other race (2)	Net effect of other race math teacher (3)	Other sex (4)	Math × other sex (5)	Net effect of other sex math teacher (6)	Joint F test (p value) (7)
A. Teacher expects ≤ high school diploma							
Full sample	0.02	0.02	0.04	0.00	0.01	0.01	0.24
[N = 16,810]	(0.02)	(0.01)**	(0.02)**	(0.01)	(0.01)	(0.01)	
White sample	−0.01	0.01	−0.00	0.00	0.01	0.01	0.49
[N = 10,600]	(0.03)	(0.04)	(0.03)	(0.01)	(0.01)	(0.01)	
Black sample	0.11	0.02	0.13	0.01	−0.01	−0.00	0.00
[N = 1840]	(0.05)**	(0.05)	(0.04)***	(0.03)	(0.04)	(0.03)	
Male sample	0.02	0.02	0.04	−0.00	0.00	−0.00	0.44
[N = 8320]	(0.02)	(0.01)	(0.02)*	(0.01)	(0.02)	(0.01)	
Female sample	0.01	0.03	0.04	−0.00	0.03	0.03	0.12
[N = 8480]	(0.02)	(0.01)**	(0.02)**	(0.01)	(0.02)*	(0.01)**	
B. Teacher expects ≥ four-year college degree							
Full sample	0.01	−0.01	−0.00	−0.01	0.01	−0.01	0.58
[N = 16,810]	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	
White sample	0.03	−0.01	0.02	−0.01	0.00	−0.00	0.98
[N = 10,600]	(0.03)	(0.04)	(0.03)	(0.01)	(0.02)	(0.01)	
Black sample	−0.07	−0.08	−0.15	−0.09	0.08	−0.02	0.06
[N = 1840]	(0.06)	(0.06)	(0.05)***	(0.04)***	(0.04)*	(0.04)	
Male sample	0.00	−0.02	−0.01	−0.00	0.02	0.01	0.23
[N = 8320]	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	
Female sample	0.01	0.00	0.01	0.00	−0.03	−0.03	0.64
[N = 8480]	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	

Notes: Each row reports estimates from a unique regression. Columns 3 and 6 report net effects of racial and sex mismatch of math teachers, respectively, which are the sum of the coefficient estimates reported in columns 1 and 2, and 4 and 5, respectively. Parentheses contain standard errors that are robust to clustering at the school level. Standard errors of the net partial effects reported in columns 3 and 6 were computed by the delta method. All models condition on student fixed effects, observed teacher characteristics, and a full set of teacher characteristic-math teacher interactions. The F tests reported in column 7 are for the joint significance of the full sets of math teacher interaction terms.
*** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

from within-student differences between two of each student's tenth grade teachers and we find no evidence of differential sorting into classrooms by race.

While our identification strategy and general interest in the consequences of student–teacher demographic mismatch are not novel (e.g., Dee, 2005), our analysis of whether teachers' expectations for specific student outcomes are biased is. This is important, as recent evidence suggests that limited information and biased expectations affect student outcomes and decision making (e.g., Hoxby & Turner, 2013; Wiswall & Zafar, 2015) and teachers likely influence students' beliefs about their academic prospects (Burgess & Greaves, 2013; Dee, 2015). Moreover, as large and nationally-representative data sets increasingly collect data on subjective expectations, assessing the role of beliefs in driving behavior will require the use of identification strategies that directly address how beliefs are endogenous or systematically biased.

Specifically, we find that non-black teachers have significantly lower educational expectations for black students than do black teachers. For example, relative to teachers of the same race and sex as the student, other-race teachers were 12 percentage points less likely to expect black students to complete a four-year college degree. Such effects were even larger for other-race *and* other-sex teachers, for black male students, and for math teachers. In addition to being statistically significant, these effects are arguably practically significant as well, as they constitute more than half of the black-white gap in teacher expectations.

The general finding of systematic biases in teachers' expectations for student attainment indicates that the topic of teacher expectations is ripe for future research. Policy relevant areas for future inquiry include how teachers form expectations, what types of interventions can eliminate biases from teacher expectations, and how teacher expectations affect the long-run student outcomes of ultimate import. The latter is a particularly important open question, as the current study's implications for educational policy and practice depend critically on the exact nature of the relationship between teachers' expectations and students' subsequent investments in human capital. For example, the direction of the effect of overly pessimistic expectations is theoretically ambiguous, as such expectations may cause students to either make ill-advised investments in higher education or motivate students to change their behaviors in ways that increase their potential and opportunities. While a thorough analysis of the impact of teachers' expectations on student outcomes is well beyond the scope of the current study, Table 10 presents some suggestive evidence that teachers' expectations do affect students' subsequent educational investments. Specifically, Table 10 reports estimates of the baseline specification (Eq. 1) in which the outcome is a binary indicator equal to one if the student subsequently enrolled in a subject-s course while in high school, and zero otherwise. For black students, and particularly for black males, having an other-race, subject-s teacher in 10th grade significantly reduces the likelihood that they enroll in another subject-s course while in high

Table 10
Linear probability model (LPM) estimates of taking a future course in same subject.

Sample	Other race (1)	Other sex (2)
Full sample	0.00	0.01
[N = 16,810]	(0.01)	(0.00)
White sample	0.02	0.01
[N = 10,600]	(0.01)	(0.01)
Black sample	−0.04	0.00
[N = 1840]	(0.01)**	(0.02)
Hispanic sample	−0.04	0.00
[N = 2110]	(0.03)	(0.01)
Male sample	0.00	0.01
[N = 8320]	(0.01)	(0.01)
Female sample	0.00	0.00
[N = 8480]	(0.01)	(0.01)
White male sample	0.03	0.00
[N = 5270]	(0.02)	(0.01)
White female sample	0.01	0.01
[N = 5330]	(0.02)	(0.01)
Black male sample	−0.08	−0.00
[N = 860]	(0.03)**	(0.02)
Black female sample	0.00	−0.01
[N = 980]	(0.02)	(0.02)
Low-income sample	0.01	−0.01
[N = 1370]	(0.03)	(0.02)
High-income sample	0.01	0.00
[N = 2770]	(0.02)	(0.01)

Notes: Each row reports coefficient estimates from the same regression. Parentheses contain standard errors that are robust to clustering at the school level. All models condition on student fixed effects and teacher characteristics. There are two observations per student, one each from the student's math and reading teacher. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

school.[16] These are precisely the groups for whom the impacts of demographic mismatch on teachers' educational expectations were largest in Table 7, suggesting that biased expectations are indeed one channel through which teachers exert long run effects on student outcomes.

To the extent that teacher expectations positively affect student outcomes, the results presented in the current study provide additional support for the hiring of a more diverse and representative teaching force, as non-white teachers are underrepresented in U.S. public schools (e.g., Kirby, Berends, & Naftel, 1999). Similarly, our results highlight the potential benefits of including expectations in teacher training and professional development program curriculums. For example, aspects of programs such as the Great Expectations (GE) initiative, which strives to ensure that all teachers nurture and help all students to reach their potential, regardless of their innate ability, talents, behaviors, or home circumstances, might be included in professional development programs nationwide (Ferguson, 2003). Finally, our results suggest that racial mismatch in other contexts in which there are asymmetries in information or status might contribute to social and economic inequities (e.g., public managers' evaluations of employees, citizens' interactions with police officers and the criminal justice system, and individuals seeking to obtain college, car, or home loans). Similar methods could be employed to study the effects of demographic mismatch on perceptions and expectations in such circumstances.

[16] This result is consistent with Fairlie et al. (2014), who find similar effects of student-instructor racial mismatch on community college students' subsequent course taking.

Acknowledgments

The authors thank Peter Hinrichs, Kyungmin Kang, an anonymous referee, seminar participants at Towson University, and attendees of the 2015 APPAM Fall Conference for offering many helpful suggestions. The usual disclaimer applies. Papageorge acknowledges that this research was supported in part by a grant from the American Educational Research Association (AERA). AERA receives funds for its "AERA Grants Program" from the National Science Foundation under NSF Grant #DRL-0941014. Opinions reflect those of the authors and do not necessarily reflect those of the granting agencies.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2016.03.002.

References

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Antecol, H., Eren, O., & Ozbeklik, S. (2015). The effect of teacher gender on student achievement in primary school. *Journal of Labor Economics, 33*(1), 63–89.

Autor, D. H., Figlio, D. N., Karbownik, K., Roth, J., & Wasserman, M. (2015). *Family disadvantage and the gender gap in behavioral and educational outcomes. Northwestern University IPR Working Paper No. WP-15-16*. Northwestern University.

Autor, D. H., Figlio, D. N., Karbownik, K., Roth, J., & Wasserman, M. (2016). *School quality and the gender gap in educational achievement. Working Paper No. 21908*. National Bureau of Economic Research.

Bailey, M. J., & Dynarski, S. M. (2011). *Gains and gaps: Changing inequality in US college entry and completion. Working Paper No. w17633*. National Bureau of Economic Research.

Boser, U., Wilhelm, M., & Hanna, R. (2014). *The power of the Pygmalion effect: Teachers expectations strongly predict college completion*. Washington, DC: Center for American Progress.

Bound, J., Turner, S., Hanushek, E., Machin, S., & Woessman, L. (2011). Dropouts and diplomas: The divergence in collegiate outcomes, *Handbook of the economics of education: Vol. 4* (pp. 573–613). Amsterdam: North-Holland.

Burgess, S., & Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics, 31*(3), 535–576.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633–2679.

Chetty, R., Hendren, N., Lin, F., Majerovitz, J., & Scuderi, B. (2016). *Childhood environment and gender gaps in adulthood. Working Paper No. 21936*. National Bureau of Economic Research.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). *How and why do teacher credentials matter for student achievement? Working Paper No. w12828*. National Bureau of Economic Research.

Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources, 48*(1), 236–264.

Curwin, R. (2012). *Believing in students: The power to make a difference*. Edutopia December 26, 2012. Accessed 31.05.15.

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.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources, 42*(3), 528–554.
Dee, T. S. (2015). Social identity and achievement gaps: Evidence from an affirmation intervention. *Journal of Research on Educational Effectiveness, 8*, 149–168.
Dillon, E., & Smith, J. A. (2013). *The determinants of mismatch between students and colleges. Working Paper No. 19286*. National Bureau of Economic Research.
Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review, 45*, 44–52.
Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial & Labor Relations Review, 48*(3), 547–561.
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), 2567–2591.
Ferguson, R. F. (2003). Teachers' perceptions and expectations and the black-white test score gap. *Urban Education, 38*(4), 460–507.
Fryer, R. G., Jr., Card, D., & Ashenfelter, O. C. (2011). Racial inequality in the 21st century: The declining significance of discrimination, *Handbook of labor economics: Vol. 4B* (pp. 855–971). Amsterdam: North Holland.
Golebiewski, D. (2012). What makes a teacher great? *Ask the students* Pittsburgh, PA: POPCity http://www.popcitymedia.com/features/kidsgreatsteachers071912.aspx Accessed 31.05.15.
Gregory, A., & Huang, F. (2013). It takes a village: The effects of 10th grade college-going expectations of students, parents, and teachers four years later. *American Journal of Community Psychology, 52*(1–2), 41–55.
Hanna, R. N., & Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy, 4*(4), 146–168.
Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources, 39*(2), 326–354.
Holt, S. B., & Gershenson, S. (2015). *The impact of teacher demographic representation on student attendance and suspensions. Discussion Paper No. 9554*. IZA.
Hoxby, C., & Turner, S. (2013). *Expanding college opportunities for high--achieving, low income students. Discussion Paper No. 12-014*. Stanford Institute for Economic Policy Research.
Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics, 27*(2), 213–256.
Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina. Working Paper No. w18624*. National Bureau of Economic Research.
Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review, 9*(2), 131–155.
Kirby, S. N., Berends, M., & Naftel, S. (1999). Supply and demand of minority teachers in Texas: Problems and prospects. *Educational Evaluation and Policy Analysis, 21*(1), 47–66.
Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics, 92*(10), 2083–2105.
Lavy, V., & Sand, E. (2015). *On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases. Working Paper No. w20909*. National Bureau of Economic Research.
Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. University of California Press.
Lareau, A., Weininger, E. B., Lareau, G., & Weininge, E. B. (2008). Class and the transition to adulthood. *Social class: How does it work* (pp. 118–151). New York, NY: Russell Sage Foundation.
Loury, G. C. (2009). *The anatomy of racial inequality*. Harvard University Press.
Lusher, L., Campbell, D., & Carrell, S. (2015). *TAs like me: Racial interactions between graduate teaching assistants and undergraduates. Working Paper No. w21568*. National Bureau of Economic Research.
McGrady, P. B., & Reynolds, J. R. (2013). Racial mismatch in the classroom: Beyond black-white differences. *Sociology of Education, 86*(1), 3–17.
Mechtenberg, L. (2009). Cheap talk in the classroom: How biased grading at school explains gender differences in achievements, career choices and wages. *The Review of Economic Studies, 76*(4), 1431–1459.
MetLife, Inc.. (2009). *The MetLife survey of the American teacher: Collaborating for student success. Part 2*. New York, NY: MetLife, Inc. https://www.metlife.com/assets/cao/contributions/foundation/american-teacher/MetLife_Teacher_Survey_2009_Part_2.pdf Accessed 05.06.15.
Morgan, S. L., Leenman, T. S., Todd, J. J., & Weeden, K. A. (2013). Occupational plans, beliefs about educational requirements, and patterns of college entry. *Sociology of Education, 86*(3), 197–217.
Mittag, N. (2012). *New methods to estimate models with large sets of fixed effects with an application to matched employer–employee data from Germany. FDZ-Methoden report*. FDZ.
Ouazad, A. (2014). Assessed by a teacher like me: Race and teacher assessments. *Education Finance & Policy, 9*(3), 334–372.
Riegle-Crumb, C., & Humphries, M. (2012). Exploring bias in math teachers' perceptions of students' ability by gender and race/ethnicity. *Gender and Society, 26*(2), 290–322.
Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review, 3*(1), 16–20.
Solon, G., Haider, S. J., & Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources, 50*(2), 301–316.
Spitzer, B., & Aronson, J. (2015). Minding and mending the gap: Social psychological interventions to reduce educational disparities. *British Journal of Educational Psychology, 85*(1), 1–18.
Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629.
Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica, 50*(6), 1415–1430.
Wiswall, M., & Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies, 82*(2), 791–824.
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT Press.