TAs Like Me: Racial Interactions between Graduate Teaching Assistants and Undergraduates*

Lester Lusher[1], Doug Campbell[2], and Scott Carrell[3,4]

[1]University of Hawaii at Manoa
[2]New Economic School
[3]University of California Davis
[4]National Bureau of Economic Research

Abstract

Over the past 40 years, institutions of higher education in the U.S. have experienced a dramatic shift in the racial composition of students enrolled in both undergraduate and graduate programs. Using administrative data from a large, diverse university in California, we identify the extent to which the academic outcomes of undergraduates are affected by the race/ethnicity of their graduate student teaching assistants (TAs). To overcome selection in course taking, we exploit the timing of TA assignments, which occur after students enroll in a course, and use within class and within student variation in TA-student race composition. Results show a positive and significant increase in course grades when students are assigned TAs of a similar race/ethnicity. These effects are largest in classes where TAs were given advanced copies of exams and when exams were not multiple choice. From a separate audit study of TA office hours and discussion sections, we find positive correlations between student and TA race, suggesting students sort according to racial match. We also find some evidence of persistent effects: Racial match improves subsequent student performance in sequenced courses, and positively influences decisions on majoring and future course enrollment for Freshmen and Sophomores. Overall, our evidence is consistent with TA-student match quality gains and role model effects.

*Special thanks to seminar participants at the University of California Davis, the Southern California Conference for Applied Microeconomists (SoCCAM), the 2016 APPAM Fall Research Conference, the AEFP 41st Annual Conference, the University of Leicester, and the University of Calgary 3rd Annual Microeconomics Workshop for helpful comments and suggestions. We also want to thank Marco for help with acquiring the data. Lastly, we want to thank the many research assistants who helped carry out the experiments: Nicholas Halliwell, Anthony Lee, Bhaverpreet Sidhu, Hongsik Woo, Catherine Jiang, Logan Schultheis, Shannon Tran, Yulia Zhestkova, Viacheslav Savitskiy, Nadezhda Kotova, and Zalina Alborova.

1 Introduction

For the past twenty years, the United States has steadily fallen relative to other developed countries in college completion rates. From 1995 to 2012, the US went from having the highest young-adult college completion rate among OECD countries to nineteenth.[1] Especially alarming is the fact that US college completion rates have stagnated *despite* increases in overall college attendance (Turner, 2004; Pew Research Center 2014) and large increases in the returns to a college education in the US (Oreopoulos and Petronijevic, 2013). Educational mobility in the US also trails the majority of other OECD countries. For example, in the US approximately 29% of men and 17% of women have less education than their parents, compared to the OECD averages of 19% and 13% of men and women, respectively (Weston, 2014). Underlining these college completion rates are prominent racial gaps. In 2015, over 50% of Asian adults aged 25 and older held a bachelor's degree or higher, compared to roughly a third of non-Hispanic Whites and less than 20% of all other races (Ryan and Bauman, 2016). Such differences in postsecondary educational attainment could lead to persistent income inequality across racial groups (Altonji and Blank, 1999; Card, 1999; Jencks and Phillips, 1998).

A natural question to ask is, once students enter college, what factors determine the likelihood they succeed and graduate? Several prior studies have presented causal evidence on various university inputs that influence undergraduate success, including capacity constraints and resources (Bound et al., 2010; Bound et al., 2012), professor quality, gender, and race (Hoffmann and Oreopoulos, 2009b; Carrell and West, 2010; Hoffmann and Oreopoulos, 2009a; Carrell et al., 2010; Fairlie et al., 2014), coaching and advising (Bettinger and Baker, 2014; Angrist et al., 2009), and academic probation (Lindo et al., 2010).

One glaring omission from this literature centers on teaching assistants (TAs), who account for nearly 15% of the total employment of postsecondary teachers in the US annually (*Bureau of Labor Statistics, Occupational Employment Statistics* 2016). TAs are graduate students employed by a university who perform various duties in the course while under the supervision of a professor or lecturer. Many of these duties impact student success in the course, including 1) hosting small weekly discussion sections, 2) holding office hours, 3) tutoring, 4) proctoring exams, 5) grading assignments and exams, and 6) arranging meetings with students. TA-student relationships are unique in that they are more likely to be a peer-based interaction, since the typical age gap between undergraduates and TAs is relatively small.[2] Additionally, with class sizes and student-professor ratios increasing in the

[1] OECD (2014), Education at a Glance 2014, Chart A3.2. Twenty-eight member countries in 2012 were considered for the study.

[2] Several studies have focused on the potential benefits of peer-based mentoring and tutoring. For example,

US (Cuseo, 2007; Kokkelenberg et al., 2008; Schanzenbach, 2014), TAs are likely to play an increasingly important role in the US post-secondary education system.

In this paper, we begin to shed light on the importance of TAs in the education production function. To do so, we focus on the role of TA race. Understanding how TA race influences student outcomes is particularly important given recent trends in the US. For the past 40 years, undergraduate and graduate programs have been experiencing a dramatic shift in student racial composition. In 1976, 82% of students enrolled in undergraduate programs in the US were White, compared to only 57% in 2013. A similar pattern can be observed in post-baccalaureate programs, where over the same time period, the fraction of Asian students has more than tripled.[3]

Why might students be influenced by TA race? Role model effects are often mentioned as an important determinant affecting educational outcomes. Another factor might include racial differences in the students' academic expectations. Research from psychology and sociology suggests that equally skilled students of different races may perform differently due to the students' self-belief about their ability to succeed, and these gaps may be muted (or exacerbated) by the TA's race (Spencer et al., 1999). Another channel is a match quality effect, where TAs of different races may have, on average, particular teaching styles or capabilities which are better suited to students of similar race.[4] Finally, TAs may exhibit bias, consciously or unconsciously, with respect to how they treat students of a similar race.

Numerous studies have investigated the importance of student-teacher interactions at the postsecondary level, with a majority focusing on the role of professor gender. Early studies found mixed results, though these studies likely suffer from potential selection biases (e.g., Rothstein, 1994; Canes and Rosen, 1995; Neumark and Gardecki, 1998). More recent studies, which have exploited within class and within student variation to overcome selection issues, have found positive same-gender effects on course grades, choice of major, course credits, and course dropping (Bettinger and Long, 2005; Hoffmann and Oreopoulos, 2009a; Hoffmann and Oreopoulos, 2009b). Likewise, using random assignment to courses, Carrell et al. (2010) find that professor gender has a significant impact on female students' performance in STEM courses. Finally, most closely related to our study, Fairlie et al. (2014) focus on student-

Castleman and Page (2014) find that near-aged peer mentors in college who sent text messages during the summer to college-intending high school graduates substantially increased subsequent college enrollment.

[3]U.S. Department of Education, National Center for Education Statistics, Higher Education General Information Survey (HEGIS), "Fall Enrollment in Colleges and Universities" surveys, 1976 and 1980; Integrated Postsecondary Education Data System (IPEDS), "Fall Enrollment Survey" (IPEDS-EF:90); and IPEDS Spring 2001 through Spring 2014, Enrollment component.

[4]This channel includes language matching where all else equal, a student learns more if particular material can be taught in the student's native language, and students who share the same race as their TAs are more likely to share the same native language.

instructor race interactions at a community college and find that race plays a large role in student outcomes.[5]

To our knowledge, this is the first study to investigate the importance of TA race. Our primary analyses utilize over 60,000 student-class observations across an eight year period (2003-2011) from a public university in California, coupled with TA assignment data from the university's Department of Economics. The institution we study is large and racially-diverse. In 2014, of the over 34,000 students enrolled, 39% were Asian or Pacific Islander, 19% were Hispanic, and 29% were White. Our data also include a survey that was offered to all professors who taught an economics class during the period of our study, which asked about exam structure (multiple choice vs. essay) and whether exams were shared with TAs prior to the exam date. Lastly, our data include an audit study conducted in 2015 which recorded student attendance during optional TA discussion sections and office hours.

We consider several empirical strategies to identify the causal effects of TA-student racial interactions and to overcome concerns of potential selection bias. Our primary analyses focus on models with class fixed effects, where we estimate differences in outcome variables between students across different races when assigned to the same TAs within the same class.[6] Since the explanatory variable varies both within class, across students, and within student, across classes, the data also allow us to control for sorting that occurs across classes by simultaneously including student fixed effects with class fixed effects. Furthermore, we find no evidence of endogenous sorting into classes by student race when predicting the race of the class' TAs with a full set of controls, including professor race and gender, student gender, high school GPA, age, class standing, and major. The lack of evidence of endogenous sorting is unsurprising for several reasons. First, the primary course registration period for undergraduates occurs before TA assignments are generated by the department. Secondly, once generated, the department only privately reveals the TA assignments to the corresponding professors and TAs.[7]

Our results show that students perform better in classes taken with TAs who are of a

[5]A handful of other studies have looked at graduate students overall as educators (e.g., Borjas, 2000; Fleisher et al., 2002; Bettinger and Long, 2004; Marvasti, 2007). Of these, Borjas (2000) and Marvasti (2007) focus on the role of nationality, using identical data sets. The authors examined student grades in response to foreign-born TAs across three introductory economics courses with an undisclosed number of teaching assistants. They find that students with international TAs received lower (self-reported) grades, and these results were particularly strong for native students.

[6]We define class as a combination of a course (e.g., Introductory Microeconomics), term (e.g., Fall 2010), and lecture. For popular courses, several lectures may be offered within the same term such that each lecture constitutes a different class. TAs are never assigned to more than one class within a term.

[7]In other words, TAs and professors themselves do not know which classes TAs are assigned to by the time most undergraduates enroll. While undergraduates still have the ability to register for courses after the end of the primary registration period, the majority of classes fill up by the time this period is over, leaving little capacity for students to be selective with their courses.

similar race. We predict a 7.7% of a standard deviation increase in course grade for students who are assigned to TAs of similar race, relative to being assigned to TAs of dissimilar race. This result is robust across various specifications, racial categorizations, and subsamples; as a falsification test, we also find that course grades are uninfluenced by the racial composition of a student's TAs in future courses. We also find some suggestive evidence of longer-run effects. Performance in the second course of a two-course sequence is significantly improved when the student was racially matched to TAs in the first course. Moreover, racial matching appears to influence student subsequent enrollment decisions, where Freshmen and Sophomores are more likely to major in Economics and enroll in another Economics class after taking a class with similarly-raced TAs. Overall, we find evidence of human capital accumulation in response to TAs.

Lastly, we introduce the audit study and professor survey to further investigate the mechanisms potentially driving the results. From the audit study, results show that students are more likely to attend their TAs' optional discussion sections and office hours when the TA is of a similar race, providing direct evidence of students responding to similarly-raced TAs. We also see that racial interaction effects are especially prominent in classes where TAs had been given copies of the exams prior to the exam dates. We interpret this result as evidence of "teaching to the exam", where TAs, perhaps unconsciously, divulge information that is pertinent to the exams. Students who attend the TAs' discussion sections and office hours are the beneficiaries of teaching to the exam, and attending students tend to be of similar race as the TAs.[8]

Racial interaction effects are also strongest in classes which had no multiple choice questions on the exams. This result could stem from several possible explanations. First, critical thinking is typically a key component to success on essay-based questions, and critical thinking skills may be fostered in settings where students discuss and ask questions about the course material, such as in TA discussion sections and office hours. Another explanation suggests that TAs are responding to students of similar race when grading.[9] Classes with no multiple choice exams are classes where TAs exercise subjective judgment when grading, and students of specific races may be more likely to answer non-multiple choice questions in a manner in which TAs of a similar race favor.[10]

[8]Given we also find improvements in subsequent course performance in two-course sequences, it could be that sharing the exam improves the teaching efficacy of the TAs such that the knowledge gained by students in the first course carries over into the second course.

[9]Other studies have found gender biases in teacher grading at the secondary school level (Lavy and Sand, 2015).

[10]Discussed in further detail later in this paper, even if grading biases were a significant channel driving the racial matching effects, studies suggest that such "grade inflation" could still lead to substantial positive human capital gains (Dee et al., 2016; Diamond and Persson, 2016). We do not believe this to be the primary channel in our setting, however, since the gains we observe in subsequent performance are restricted

The remainder of this paper proceeds as follows. Section 2 introduces the data. Section 3 discusses our identification strategies and econometric specifications. Section 4 presents our results. Section 5 discusses the potential underlying mechanisms of our results, and section 6 concludes.

2 Data

2.1 Data Sources and Institutional Background

Our paper centers on detailed student administrative data from a large, public university in California with a highly diversified student community. In 2014, over 34,000 students enrolled at the university, where 39% of the enrolled students were Asian or Pacific Islander, 19% were Hispanic, and 29% were White. U.S. News & World Report (2015) classifies the university admissions as "most selective" and ranks the university as one of the best public university in the United States. Our primary analyses link the student administrative data to graduate teaching assistant (TA) assignment data from the university's Economics Department. The B.A. in Economics is the second largest major at the university, accounting for over 6% of degrees conferred annually. These data cover the academic school years from 2003 to 2011 for the three primary quarters of enrollment: Fall, Winter, and Spring.[11]

Each observation in our primary data set pertains to a student who enrolls in an Economics class. We define a class as a combination of a course (e.g., ECN100), a term (e.g., Fall 2010), and a "lecture". Every term, the department offers a series of courses. These course-terms typically constitute a single "class" which has a single syllabus, professor, and up to three TAs. For very popular courses, sometimes the department will offer multiple "lectures" within a single term; though they share the same course-term, these lectures are very different classes, as they each contain their own syllabus, meeting times, exams, TA(s), and (almost always) professor.[12] We have a series of student-level characteristics, including term admitted, admission basis (freshman vs. transfer), gender, race, nationality, parental education, and high school GPA. Student-by-term level variables include academic standing (Freshman/Sophomore/Junior/Senior), major(s) when they registered for the course, and age. Finally, class-level controls include professor gender and race.

We match each student by class observation to TAs assigned to the class. Since a single class may contain up to three TAs, we do not necessarily analyze one-to-one matches between

to two-course sequences, where knowledge spillovers are especially pertinent.

[11]Hence, we do not focus on any special quarters, such as summer sessions.

[12]For instance, Principles Microeconomics for the Fall 2011 term could have had two different "classes", where the first class was taught by Professor Xavier with TAs Scott, Logan, and Jean, while the second class was taught by Professor Oak with TAs Ashe and Brock. To make a distinction between these two course-terms, we say that they have different "lectures". These two classes each have different professors, TAs, meeting times, exams, etc.

a student and a TA. In Economics courses, a student is technically assigned to a single TA, but often has the liberty to choose any of the TAs in their class to attend discussion sections, visit office hours, etc. Furthermore, TAs within a class jointly share numerous responsibilities, including assisting with lectures and grading assignments and exams. Consequently, as described in further detail in section 3, we link the race of a student enrolled in a class to the racial composition of the TAs assigned to the class.

Lastly, our paper utilizes two supplemental sets of data. First, in the Fall of 2014, a survey was offered to all professors who taught a class during our 2003 to 2011 time frame. For each class a professor taught, the survey recorded 1) whether the professor shared a copy of the class' exams with the TAs prior to the exam date and 2) the structure of the exams (multiple choice vs. short/long answer).[13] Approximately 58% of our total student-by-class observations are covered by professor survey responses. Secondly, in the Spring 2015 quarter, an audit study was conducted where student attendance by gender and broad racial categories at TA discussion sections and office hours was recorded by an undergraduate research assistant who audited the class. TA discussion sections and office hours are hosted weekly throughout the quarter, and attendance in this setting is optional for enrolled students. Auditors visited the TA discussion sections during the third and fourth weeks of the term and the office hours during the fifth and sixth weeks. The audit study covers 124 discussion sections and 102 office hours.

2.2 Summary Statistics

The main outcome variable of interest is the grade each student received in each class, conditional on staying enrolled. Following the classical American letter grading system, at the end of the term, each professor assigns a letter grade (with +/- modifiers) to each student in his/her classes based on the student's performance on class assignments and exams. Each letter grade then gets translated by the university into a numerical grade point average (GPA) value (e.g., $A = \frac{12}{3} = 4.0$, $A- = \frac{11}{3} \approx 3.7$, $F = 0$). For each class, professors are asked to attain an average GPA around 2.7, though professors are given discretion to deviate from this average. For our primary analyses, we standardize each student-by-class grade to a mean of zero and a standard deviation of one by class, and call this variable "Standardized grade". Another outcome of interest is "Passed class", which is an indicator that switches on if the student received a C- ($\frac{5}{3}$ in GPA units) or higher, or a "Pass" in the class.[14] Later

[13] Valid responses to the question of sharing exams with the TAs included "Yes", "No", "Sometimes", and "I don't remember". When using this question for analyses, we only focus on the sample of responses that were either "Yes" or "No".

[14] At some point toward the beginning of the term, students can opt to receive a grade of either "Pass" (P) or "No Pass" (NP) on their transcript for their class, even though the class is otherwise graded using the standard letter grading system. At the end of each term, instructors submit the letter grades received

analysis also considers ordered logit and ordered probit specifications using the raw letter grades as outcomes.

Table 1 presents summary statistics for our main sample of interest. We have 60,642 student-by-class observations, 19,522 students, 614 classes, and 286 teaching assistants. From Panel A, over 50% of students were male, and nearly 4% of students are identified as international, while the average high school (weighted) GPA for students was 3.64. Panel B reveals that on average, over 117 students enrolled in each class, over 70% of instructors were White, and nearly 14% of instructors were Asian. Consistent with the notion that classes are difficult to get into, Panel C shows that under 2% of students drop the class once they successfully enroll. "Numeric course grade" in Panel C corresponds to the numerical grade the student received in the class, which ranges between 0 (for F) and 4 (for A). White students received an average class grade around 2.55, in between a C+ and B-, while Asian students received an average grade slightly above a B-. The standard deviation of GPA is around one grade point unit, so the distance between two letter grades (e.g., C vs. B) is roughly one standard deviation. Over 80% of students passed their classes. From Panel D, we find that almost half of the students were Asian, while nearly a third were White. Meanwhile, for TAs, roughly 38% were White and 44% were Asian.[15]

Given these summary statistics, we can also turn to the Integrated Postsecondary Education Data System (IPEDS) to consider how generalizable our findings may be to other universities. The fraction of White students in our sample compares similarly to that of all other California universities, where 27.1% and 36.9% of undergraduate and graduate students were White, respectively. On the other hand, our sample has a smaller share of minorities and a larger share of Asian students, as roughly 22% of California undergraduates were Asian and 50% were from a minority background. For other observable characteristics, undergraduates in our sample appear similar to the average undergraduate attending a US research university.[16] For instance, the 25^{th} (75^{th}) percentile SAT Math score for undergraduates at research universities was 558 (667), compared to 560 (680) for our sample. Our sample contains 47% female students, compared to 51% of undergraduates at research universities. Approximately 5% of students from research universities come from a foreign

by each student in the class to the registrar's office, and for students who switched into P/NP grading, the registrar converts all grades above and including C- to P, while grades below C- get converted to a NP. For this subset of students, we can only observe their final P or NP grade, and not the letter grade they received prior to the conversion.

[15] A student/TA/professor is classified as of Asian race if their primary race is recorded as Chinese, Japanese, Korean, Filippino, South-East Asian, Vietnamese, Thai, or "Other Asian". TA race was collected by utilizing a combination the TAs' names, information on the TAs' personal websites (e.g., pictures, CVs), and intermediaries' personal knowledge of the TAs.

[16] IPEDS uses the basic Carnegie Classification of Institutions of Higher Education (2010) to determine the list of research universities.

country, and the age profile of undergraduates also looks broadly similar.

3 Econometric Specifications

Our primary analysis estimates the following specification:

$$y_{ikt} = \psi(Asian_i * AsianTA_{ikt}) + \beta X_{it} + \lambda_{kt} + \alpha_{kA} + \delta_{tA} + u_{ikt} \quad (1)$$

where y_{ikt} is an outcome for student i taking course k in school term t, $Asian_i$ is an indicator variable for whether student i is Asian, $AsianTA_{ikt}$ is the fraction of student i's TAs for class kt that were Asian, X_{it} is a vector of student-by-term controls, and λ_{kt}, α_{kA}, and δ_{tA} are class, course-by-race, and term-by-race fixed effects, respectively. Since the number of TAs assigned to a class ranges from one to three, $AsianTA_{ikt}$ carries a value of either 0, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$, or 1.

The core of our identification strategy centers on class fixed effects λ_{kt}, which control for unobserved factors that vary at the class level and affect student performance. Note that class fixed effects also control for professor fixed effects since each class is taught by exactly one professor. These, in turn, control for the possibility that students of a particular race take classes with professors who are systematically different from other professors. Class fixed effects also avoid the need to rely on settings with standardized grading or testing procedures across classes since students within a class are completing the exact same assignments and tests. Thus, we are solely comparing the academic performances of Asian and non-Asian students within the same class and subjecting the students to the same class-level shocks, such as the professor's and TAs' characteristics (e.g., ability/experience) or the time/size of the class. Course-by-race fixed effects allow for racial differences in the outcome variable to vary across courses. These are necessary to account for the possibility that the courses in which non-Asians and Asians tend to perform differently are also the courses in which TAs tend to be non-Asian or Asian, respectively.[17] Term-by-race fixed effects account for the possibility that the academic capabilities of Asian or non-Asian students are changing over time. The coefficient ψ measures the average outcome gain for Asian students, relative to non-Asian students, from assignment to Asian TAs. Conversely, ψ measures the average outcome loss for non-Asian students, relative to Asian students, from assignment to Asian TAs versus non-Asian TAs.

To measure student attendance-by-race to TA discussion sections and office hours from

[17]For example, Asian students may be more likely to enroll in an international studies course and Asian TAs may be more likely to be assigned to international studies. Indeed, when evaluating student grades, our estimated magnitude of ψ slightly increases when we exclude course–by-race fixed effects (see Appendix Table A.7).

the audit study, we consider the following specification:

$$fracStudentAsian_s = \rho AsianTA_s + \beta X_s + u_s \quad (2)$$

where each observation corresponds to TAs' discussion sections or office hours. X_s comprises of indicators for the weekday, the time, and the individual auditor for the discussion section or office hour.[18] Observations are weighted by total attendance of students to the discussion section or office hour. The coefficient ρ is the expected increase in the fraction of attendees who are Asian in response to the discussion section or office hour being hosted by an Asian TA.

3.1 Identification

The primary threat to our identification strategy is self-selection into courses by TA race, which could result in a correlation between unobserved variables in the error term u_{ikt} and the interaction term $Asian_i * AsianTA_{ikt}$. For example, our estimates would be biased if high ability Asian students systemically select into classes assigned Asian TAs and high ability non-Asian students systematically select into classes assigned non-Asian TAs. Prior work looking at professor-student relationships potentially suffer from such selection biases, where students of a particular gender/race, and different academic capabilities, select into classes based on the teacher gender/race.[19]

To mitigate selection biases, previous studies have often focused on a sub-sample of students or classes where selection was arguably less of an issue.[20] Fortunately, in our setting, it is nearly impossible for undergraduates to identify which TAs are assigned to classes prior to enrollment. Importantly, the primary registration period for undergraduate classes occurs well before the economics department generates TA assignments for classes.[21] While undergraduates (technically) have the ability to register for courses after the end of the primary registration period, the majority of classes fill up by the time this period is over, leaving little capacity for students to be selective with course registration.

To formally test for endogenous enrollment by race, we collapse our data to the class level, and regress the racial composition of the class TAs on observable characteristics and term and course fixed effects. The first and second columns of Table 2 presents results

[18] There were 23 separate auditors who attended the discussion sections and office hours.

[19] Perhaps exacerbating selection biases in prior studies are services such as *ratemyprofessor.com*, which provide students with extensive information about their instructors.

[20] For instance, Fairlie et al. (2014) focus on students with relatively low standing on registration priority lists since these students have little ability to be selective with their courses.

[21] For example, for the Spring 2014 term, which started in March, the primary undergraduate registration period started on February 3 and ended on February 14. The Economics department generated and privately revealed TA assignments on February 27 to TAs and professors.

from our main sample of interest, while the next two columns consider the subsample of classes taught by professors who completed our survey. The odd columns consider all our primary regressors, while the even columns include interactions between student race and the primary regressors. Results show that our regressors are generally small and weak predictors of Asian TA composition, and the race of the student is a weak predictor of TA race. For each regression, we test the hypothesis that all covariates are jointly equal to zero, conditional on term and course fixed effects, and report the p-values. Across all samples, we fail to reject the hypothesis that all covariates have no power in predicting TA race.[22] Results from this analysis, coupled with practical knowledge of the registration process for students into classes, indicate that our primary regressor of interest is likely free from selection bias.

4 Results

4.1 Main Results

Table 3 presents our main results. In the first row, we report the coefficient on the Asian TA racial composition, while the second row "Effect of Similar Race" reports the estimated coefficient of interest ψ. Recall that ψ can be interpreted as the expected relative change in performance between Asian and non-Asian students when the student has all Asian TAs instead of all non-Asian TAs. From column one, the -0.037 estimate, together with the $Asian_i * AsianTA_{ikt}$ control, suggests that non-Asian students receive a 3.7% of a standard deviation decrease in their grade in response to enrollment with Asian TAs; conversely, non-Asian students receive a 3.7% standard deviation of a grade boost in response to non-Asian TAs. Moreover, since the sum of the two coefficients is 0.024, we can see that Asian students receive a 2.4% of a standard deviation increase in grade in response to assignment to Asian TAs. The positive coefficient for $Asian_i * AsianTA_{ikt}$ implies students do relatively better when matched to TAs of similar race.[23]

For the remaining columns of Table 3 under "Standardized grade", we consider the sensitivity of the results to the inclusion of different fixed effects and controls. The second column

[22] We also consider additional tests of endogenous enrollment in the Appendix. First, we mimic the "sorting regressions" of Fairlie et al. (2014) in Table A.1, and find no evidence of endogenous sorting. The primary benefit of the Fairlie et al. (2014) specification is the ability to condition on class fixed effects. A drawback is that one cannot simultaneously test the importance of observables X_{ikt} in predicting TA race. Second, in Table A.2, we consider the same test in the main text but only using characteristics of students who completed the course; these results are very similar to those in Table 2, highlighting how drop rates are low in our setting. Lastly, we consider our tests for endogenous enrollment using the student-class level data in Table A.3, and again we find little evidence of sorting.

[23] A positive racial interaction effect could still have had arisen if, for example, all students irrespective of race did better with Asian TAs, but Asian students performed especially well relative to non-Asian students when matched to Asian TAs. If this were the case, then our estimated coefficient on $AsianTA_{ikt}$ would have been positive, and the model would predict a drop in grades for non-Asian students in response to non-Asian TAs.

includes a full set of controls, while the third column replaces course and term fixed effects with class fixed effects.[24] Column four replaces student-level controls with student fixed effects.[25] Across specifications, we estimate statistically significant gains for students when assigned TAs of similar race. From our fully-specified model with both class and student fixed effects, we predict a 0.077 standard deviation increase in course grade when students are matched to TAs who are all of a similar race as themselves.[26] Given the standard deviation of course grades is slightly over one grade point unit, and the value of a grade modifier (+/-) is a third of a grade point unit, this effect is roughly equivalent to an increase of a fourth of a grade modifier. Standard errors are clustered by professor for all specifications.[27] Panel B presents the results for classes where the professor responded to our survey.[28]

The last four columns of Table 3 consider an indicator for whether the student passed the class as an outcome variables. Recall that "Passed class" is essentially just an indicator for whether the student received a letter grade of C- or higher. Consistent with the results for standardized grade, we again find positive racial interaction effects: both Asian and non-Asian students experience increases in the likelihood of passing the class when assigned to Asian and non-Asian TAs, respectively. For our fully-specified model with class and student fixed effects, we estimate a positive racial interaction effect of three percentage points, which is statistically significant at the 1% level.[29]

[24] Note that the coefficient on TA race is not estimable with class fixed effects due to collinearity.

[25] Since student fixed effects rely on within student, across class variation, our primary regressors of interest are only identified with students who enrolled in more than one class.

[26] When the data are parsed by White and non-White students and TAs, this coefficient drops slightly to 0.076 standard deviations, and maintains statistical significance at the 1% level (see Appendix Table A.4). Racial interactions remain statistically significant when we consider specifications with finer race categorizations (see Appendix Table A.5 and Table A.6).

[27] With fewer professor clusters than class clusters, we conservatively cluster at the professor level instead of the class level. Ideally, we would cluster at the TA level, but since a single class may contain up to three TAs, a single observation may belong to up to three TA clusters. There are fewer professors than TAs in our setting. As a robustness check, we consider the subsample of classes which had only one TA and cluster at the TA level. Both estimated magnitudes and standard errors slightly increase, with the results remaining largely statistically significant (see Appendix Table A.7, Panel C). In general, standard errors decrease when we cluster at the class level instead of professor level.

[28] Panels A and B of Appendix Table A.7 investigate the sensitivity of the results to further combinations of controls and fixed effects for the main sample and professor survey sample, respectively.

[29] One hypothesis for these results is that the racial interaction effects simply reflect a systematic change in student composition that occurs after students enroll in the class. That is, after students observe the race of their TAs, they decide whether to drop the course. While this effect is likely to be small, since overall drop rates are under 2% (Table 1), we formally test for this possibility using "Dropped class" as an outcome variable in Appendix Table A.8. We find a very small, positive, and statistically insignificant racial interaction effects, suggesting, if anything, that students are slightly *more* likely to drop a class with TAs of similar race.

4.2 Specifications Using Letter Grades

In order to test the robustness of our results further, and to understand how the distribution of grades shifts in response to TA race, we consider alternative specifications utilizing the raw letter grades students received in their classes. Students who enrolled in a class for a letter grade received either an A(-), B(+/-), C(+/-), D(+/-), or F. Figure 1 displays the marginal effects from ordered logit regressions on the probability of attaining each possible letter grade. Similar to our main specification, "Effect of Similar Race" reports the coefficients on $Asian_i * AsianTA_{ikt}$, while the remaining two sets of estimates report the marginal effects for Asian/non-Asian students in response to Asian/non-Asian TA racial composition, respectively.[30] We find that students are significantly more likely to attain grades of B or higher when matched to TAs of similar race. Correspondingly, students are also less likely to attain grades of C+ or lower when matched to TAs of similar race. The largest marginal effects come from increases in the probability of receiving an A, followed by decreases in the probability of receiving a C, in response to having TAs of similar race.

4.3 Audit Study of TA Section and Office Hour Attendance

Next, we turn to the Spring 2015 audit study to test for student response to TA race by examining student attendance at optional TA discussion sections and office hours. Results in Table 4 show that across all variations of specification (2) and samples, TA race is positively related to the race of the attending students. From column three, we predict an 8.4 percentage point increase in the fraction of attending students who are Asian in response to the discussion section being taught by an Asian TA. For office hours, we estimate a 20 percentage point increase in fraction of Asian attendees in response to the office hour being hosted by an Asian TA.[31] Given the lack of association between student race and TA race in student enrollment from Table 2, correlations in race at TA discussion sections and office hours can only occur if within each class, students are sorting across discussion sections and office hours to match their TA's race.[32] More precisely, this would occur if a student has multiple TAs and they choose to visit the TA with similar race, or if the student has no choice in TA race, then they decide whether just to skip discussion sections or office hours.

[30] Appendix Table A.9 reports the estimated coefficients and marginal effects from these regressions, as well as estimates from ordered probit models. Estimates across ordered logit and ordered probit specifications are nearly identical and exhibit similar patterns.

[31] The total number of (non-)Asian students who attend discussion sections increases when the TA is (non-) Asian. On the other hand, the total number of office hour attendees decreases, irrespective of student race, in response to office hours being hosted by an Asian TA; the decrease in Asian student attendees is smaller than the decrease from non-Asian student attendees.

[32] This is assuming enrollment patterns in Spring 2015 reflect the lack of endogenous enrollment from our main results, which utilize data from 2003 to 2011 and condition on term and course fixed effects.

4.4 Professor Survey Samples

We now turn to the professor survey portion of our main sample to examine whether our racial interaction effects differ across classes based on the professors' responses. These results are presented in Table 5, where we consider our main specification with class fixed effects and a full set of controls.[33] Responses to both of the two questions on the professor survey appear to determine which classes are driving the racial interaction effects. First, we find that the effects are particularly driven by classes where TAs were given advanced copies of the exam, where we estimate a racial interaction effect of 0.127; meanwhile, in classes where the exams were not shared with TAs, we estimate a statistically insignificant -0.003 racial interaction effect. We also find that the effects are largest in classes with non-multiple choice exams (0.199), while we observe smaller, statistically insignificant interaction effects when focusing on classes that had exams with multiple choice (0.046).[34]

We interpret these first results as TAs "teaching to the test", where when a TA is given a copy of the exam, the TA adjusts his/her discussion section and office hour lessons to better suit the material that will appear on the exam. Teaching to the test would benefit students who attend discussion sections and office hours, and as suggested by the audit study, attending students tend to be of similar race as the TA. We posit that the latter results by classes with multiple choice exams could be driven by several explanations. First, it could be that classes without multiple choice exams may require more critical thinking skills, which are gained in TA discussion sections and office hours. Another explanation stems from TA grading behavior, since classes with non-multiple choice exams allow TAs to exercise more subjective judgments when grading. Additionally, students of specific races may be more likely to answer non-multiple choice questions in a manner which TAs of a similar race favor. We return to these results in section 5 when we further discuss potential mechanisms underlining the racial interaction effects in conjunction with results in the next two sections.

4.5 Specifications with Past and Future TA Race Interactions

In this section, we consider our main specification while including an additional regressor of interest: student race interacted with the student's TA racial composition from previous or future courses. Examining if student performance in current courses responds to TA race in *future* courses serves as an additional validity check of our identification strategy. Under the assumption that there are no across-term correlations of selection into courses by TA race, the race of a student's TAs in future courses should have no influence on a student's current

[33] Table A.10 presents professor survey results across a variety of alternate specifications.

[34] We also find that the racial interaction effects are strongest in classes that had both shared exams with TAs and no multiple choice (see Appendix Table A.11).

term performance. Panel A of Table 6 presents these results, where the additional regressor is $(Asian_i \times AsianTA_{i(t+1)})$. Across all specifications, we find that the race of future TAs has no impact on current term performance.[35]

Regressions where we interact student race with the student's past TA race test for potential spillover effects. If material across courses had significant overlap and there were significant increases in learning in response to TA race, then the racial composition of a student's *past* TAs would influence their current grades. To test for this possibility, we again consider our main specification where we additionally include an interaction term between the indicator for student race $Asian_i$ and the proportion of student i's TAs who were Asian in the term prior to the class being taken. Panel B of Table 6 presents results from this analysis. Under these specifications, we find no evidence that the racial composition of a student's prior TAs influences current grades.

Finally, in Panel C of Table 6, we again consider an interaction between student race and past TA race, but instead focus strictly on the full set of "sequenced" courses in the data. In the Economics department in our setting, a total of nine two-course sequences exist, where either the first course is a prerequisite for the second course, or the courses come from the same subfield within Economics.[36] These are courses for which a test for positive spillovers is likely more appropriate, particularly since material across these two courses should have significant overlap. The additional regressor included is an interaction between student race and the TA racial composition from the student's first course in the two-course sequence ($AsianTA_{prev.course}$). This analysis suffers from significantly reduced statistical power, particularly since only 7,430 of our student-class observations come from enrollment in the second course of a two-course sequence (where students also enrolled in the first course). These results do provide some evidence of positive spillovers. Though we only attain statistical significance at the 10% level in three of seven specifications, we find lagged interaction effects ranging between 0.081 and 0.118. In total, and discussed in further detail later in the paper, we interpret these results as suggestive evidence of increased learning and human capital accumulation for students when matched to TAs of similar race.

[35] Future race is calculated by taking the average TA racial composition across a student's classes in the subsequent term of classes enrolled. Note that this analysis drops all students who only enrolled in one term of classes, and the last term for which students enrolled in classes.

[36] These two-course sequences include (Introductory/Intermediate) Microeconomics, (Introductory/Intermediate) Macroeconomics, (Introductory/Intermediate) Econometrics, Industrial Organization (A/B), Labor Economics (A/B), World Economic History (A/B), US Economic History (A/B), International Economics (A/B), and Public Economics (A/B).

4.6 Subsequent Course Enrollment and Major Choice

Lastly, we return to our main specification, with class fixed effects and our full set of controls, to explore whether there are persistent racial matching effects in the form of student decisions to enroll in more Economics courses and to major in Economics. We define "Enroll in Another Class" as a student-class level indicator for whether the student enrolled in another Economics class in a later term, and "Major in Economics" as a student-level variable that indicates whether the student was declared as an Economics major during their last Economics class at the university.[37]

Table 7 presents these results. For our full sample, we find positive, but statistically insignificant, racial interaction effects on course enrollment (0.008) and majoring (0.007). When we focus on various subsamples, we start to see more convincing evidence of persistent effects. For instance, Freshmen and Sophomores are 2.3 percentage points more likely to enroll in another Economics class in response to racial match, compared to just 0.1 and 0.4 percentage points for Juniors and Seniors, respectively. Freshmen and Sophomores are also statistically significantly more likely to major in Economics in response to racial match. The only two subsamples for which we detect statistical significance for subsequent enrollment is for classes where exams were shared with TAs, and where exams were not multiple choice; this result perfectly mirrors the observed boosts in grades in the professor survey sample. Though statistically insignificant, we also find non-Economics majors, relative to Economics majors, are more likely to enroll in another Economics class in response to racial match. Relative to students in advanced classes, students in introductory Economics courses are more likely to major in Economics in response to racial match. Finally, we find some evidence that the racial interaction effects had particularly positive persistent effects on first generation college-going students: these students were 2.3 percentage points more likely to enroll in another Economics class and 2.8 percentage points more likely to major in Economics after taking a class with TAs of similar race.

5 Discussion of Potential Mechanisms

An important question to address for welfare and potential policy implications centers on the mechanisms that are driving our results. TA race could influence student outcomes in several manners. Role model effects are often mentioned as a determinant affecting educational

[37]Unfortunately, we can only observe a student's declared major at the time they enrolled in the course. Thus, we cannot observe any instance where a student switched their major after their final Economics course at the university. For example, any incoming students declared as Economics majors who subsequently enrolled in an introductory Economics class and experienced a negative outcome with dissimilarly raced TAs may have decided to switch out of the Economics major and refrain from enrolling in another Economics course. To account for this, we coded any students who were declared as an Economics major as not finishing with an Economics major if they did not enroll in any classes beyond the two introductory courses.

outcomes. In our setting, students may be inspired by their TAs, or be more comfortable approaching and learning from their TAs due to the TA sharing a similar race.

Another channel is a match quality effect, where TAs of different race/ethnicity have particular teaching styles which are better suited to students of similar race/ethnicity. Included in this channel is a language matching effect, where students learn more if course material can be explained in the student's native language, which is more likely to occur when students share the same race/ethnicity as their TAs. Thus, with a match quality effect, students and TAs are not directly responding to the other's race, but instead students are reacting to a characteristic that is, on average, associated with their TAs' race/ethnicity.

Finally, TAs could exhibit bias with respect to how they treat students of a similar race. Discrimination could happen on an unconscious level where, for example, TAs of particular races may be more lenient when grading certain types of errors on exams that are more likely to be made by students of similar race. Such "grade inflation" could still lead to subsequent learning and human capital gains through a type of "self-signaling" model (e.g., Dee et al., 2016; Diamond and Persson, 2016).

Though we cannot rule out potential TA biases, we believe our results to be most consistent with role model effects and match quality effects. To start, the audit study provides direct evidence of students responding to the TA race in the form of voluntary attendance. Furthermore, the underlying motive for the students' attendance is likely driven by a match quality effect, where a student is learning more from their TAs due to the TAs' teaching styles or capabilities. We also find the racial matching effects to be strongest in classes where exams were shared with TAs; we believe this to be evidence of "teaching to the test", and students of similar race as the TA particularly benefit from the TA "teaching to the test" since similarly raced students are more likely to attend the TA's discussion sections and office hours. The strongest piece of evidence in potential TA biases lies in the stronger effects in classes without multiple choice exams; it could be that students of specific races may be more likely to answer non-multiple choice questions in a manner which TAs of a similar race favor.[38] Of course, it could also be that classes that rely less on multiple choice exams are classes where student-TA relationships are generally more important for learning the material and succeeding in the class.

Lastly, the persistent effects from racial matching are perhaps the most convincing pieces of evidence of significant learning and human capital gains driven by role model effects and match quality gains. The only feasible channel through which grading biases could lead to a persistent student response is discussed and identified in recent work from Diamond

[38] For example, perhaps Asian TAs are more likely to be forgiving of an answer written with poor grammar, and perhaps Asian students are more likely to write with poor grammar.

and Persson (2016) and Dee et al. (2016); these papers find that "grade inflation" (giving higher scores than those earned based on performance) led to boosts in subsequent student performance. In our setting, it could be that students' grades are inflated in response to TA bias, which then serves as a type of self-signal for the student, which subsequently positively influences their decisions about enrollment and effort in future Economics courses. We believe, however, that this channel is unlikely to apply to our setting, namely because we only find increased subsequent performance for "sequenced courses", or courses which had significant overlapping material. To argue a self-signaling model, Diamond and Persson (2016) show that test score inflation led to boosts in subsequent performance in *other, unrelated* subjects. In our setting, if one were to assume that the finest level with which grade inflation boosts a student's belief is in Economics overall, then we should see improvements in subsequent performance across all Economics courses, which is not the case in our setting (Panels A and B of Table 6). In total, it appears biases in grading could, at most, explain only a small part of the effects, while role model effects and match quality gains are likely the biggest drivers.

6 Conclusions

In spite of increases in overall attendance, college completion rates have stagnated in the US. A natural question to ask is, once the student enters college, what factors determine student success? The goal of this paper is to shed light on the importance of TAs in determining student outcomes, focusing on the role of TA race. Understanding how TA race influences student outcomes is especially important given recent trends in the US, where the racial composition of undergraduate and graduate programs have been experiencing dramatic shifts over the past 40 years.

Our primary analyses come from detailed student administrative data from a large public university in California, paired with TA assignment data from the university's department of Economics. We consider several empirical strategies to overcome concerns of potential selection bias. We first focus on models with class fixed effects, where we compare differences in outcomes between students across different races when assigned to the same TAs within the same class. Furthermore, we simultaneously control for sorting that occurs across classes by including student fixed effects. We find no evidence of endogenous sorting into classes by student race when predicting the race of the class' TAs with a full set of controls. The lack of sorting is unsurprising since students have very little ability identifying which classes TAs are assigned to, and TA assignments are generated after the undergraduates' primary registration period ends.

We find that students perform better when taking a class with TAs who are of a similar race. Students are more likely to attend their TA's optional office hours and discussion sections when the TA is of a similar race. Racial interactions are strongest in classes where TAs had been given a copy of the exam prior to the exam date, and when the exams for the class had no multiple choice. We also find evidence of persistent effects in the form of subsequent course performance, course enrollment, and majoring choice. Performance in the second course of a two-course sequence is significantly improved when the student was racially matched to TAs in the first course. Lastly, Freshmen and Sophomores are more likely to enroll in another Economics class, and major in Economics, after taking a class with similarly raced TAs. Overall, we find evidence of student learning and human capital accumulation in response to TAs that can primarily be attributed to role model effects (students responding to TAs of similar race) and match quality gains (TAs teaching style or capabilities better match students of similar race).

Even if the observed racial matching effects were strictly driven by increases in student effort and learning, the policy implications are still not completely straightforward. While it is not the case in our specific setting, student-TA racial matching effects could still exist in a setting where all students, irrespective of race, perform better in response to a particular race. For instance, it could be that all students see boosts in performance in response to having White TAs, but that the boost is larger for White students. Hence, one would face an efficiency-equity tradeoff if one were to consider sorting of TAs/students to maximize racial matching; namely, students who share a race with TAs of overall lower quality would miss out. If it was instead the case that all students see boosts in performance when matched to similarly raced TAs (as is the case in our setting), then the department/university could consider sorting TAs/students to maximize racial match and improve performance for all students. Finally, another potential policy margin of interest is on the employment of TAs. While it could just be viewed as an "extensive" margin of sorting, departments could consider increasing the employment of TAs to better match the racial composition of the student body, thus increasing the overall likelihood students racially match their TAs.

References

Altonji, J. G. and R. M. Blank (1999). "Race and Gender in the Labor Market." In: *Handbook of Labor Economics* 3, pp. 3143–3259.

Angrist, J., D. Lang and P. Oreopoulos (2009). "Incentives and Services for College Achievement: Evidence from a Randomized Trial." In: *American Economic Journal: Applied Economics*, pp. 136–163.

Bettinger, E. P. and R. B. Baker (2014). "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising." In: *Educational Evaluation and Policy Analysis* 36.1, pp. 3–19.

Bettinger, E. P. and B. T. Long (2005). "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students." In: *American Economic Review*, pp. 152–157.

Bettinger, E. and B. T. Long (2004). *Do college instructors matter? The effects of adjuncts and graduate assistants on students' interests and success*. Tech. rep. National Bureau of Economic Research.

Borjas, G. J. (2000). *Foreign-born Teaching Assistants and the Academic Performance of Undergraduates*. Tech. rep. National Bureau of Economic Research.

Bound, J., M. F. Lovenheim and S. Turner (2010). "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." In: *American Economic Journal: Applied Economics* 2.3, pp. 129–57.

Bound, J., M. F. Lovenheim and S. Turner (2012). "Increasing Time to Baccalaureate Degree in the United States." In: *Education Finance and Policy* 7.4, pp. 375–424.

Bureau of Labor Statistics, Occupational Employment Statistics (2016).

Canes, B. J. and H. S. Rosen (1995). "Following in Her Footsteps? Faculty Gender Composition and Women's Choices of College Majors." In: *Industrial and Labor Relations Review*, pp. 486–504.

Card, D. (1999). "The Causal Effect of Education on Earnings." In: *Handbook of Labor Economics* 3, pp. 1801–1863.

Carrell, S. E., M. E. Page and J. E. West (2010). "Sex and Science: How Professor Gender Perpetuates the Gender Gap." In: *The Quarterly Journal of Economics* 125.3, pp. 1101–1144.

Carrell, S. E. and J. E. West (2010). "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." English. In: *Journal of Political Economy* 118.3,

Carrington, B., P. Tymms and C. Merrell (2005). "Forget Gender: Whether a Teacher is Male or Female Doesn't Matter." In: *Teacher: The National Education Magazine* Dec 2005, p. 32.

Carrington, B., P. Tymms and C. Merrell (2008). "Role Models, School Improvement and the 'Gender Gap'—Do Men Bring out the Best in Boys and Women the Best in Girls?" In: *British Educational Research Journal* 34.3, pp. 315–327.

Castleman, B. L. and L. C. Page (2014). "Summer Nudging: Can Personalized Text Messages and Peer Mentor Outreach Increase College Going Among Low-income High School Graduates?" In: *Journal of Economic Behavior & Organization*.

Cuseo, J. (2007). "The Empirical Case Against Large Class Size: Adverse Effects on the Teaching, Learning, and Retention of First-Year Students." In: *The Journal of Faculty Development* 21.1, pp. 5–21.

Dee, T. S. (2004). "Teachers, Race, and Student Achievement in a Randomized Experiment." In: *Review of Economics and Statistics* 86.1, pp. 195–210.

Dee, T. S. (2005). "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" In: *American Economic Review*, pp. 158–165.

Dee, T. S. (2007). "Teachers and the Gender Gaps in Student Achievement." In: *Journal of Human Resources* 42.3, pp. 528–554.

Dee, T. S. et al. (2016). *The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations*. Tech. rep. National Bureau of Economic Research.

Diamond, R. and P. Persson (2016). *The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests*. Tech. rep. National Bureau of Economic Research.

Ehrenberg, R. G., D. Brewer and D. Goldhaber (1995). "Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from the NELS." In: *Industrial and Labor Relations Review* 48.3, pp. 547–561.

Fairlie, R. W., F. Hoffmann and P. Oreopoulos (2014). "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." In: *American Economic Review* 104.8, pp. 2567–91.

Fleisher, B., M. Hashimoto and B. A. Weinberg (2002). "Foreign GTAs can be effective teachers of economics." In: *The Journal of Economic Education* 33.4, pp. 299–325.

Hoffmann, F. and P. Oreopoulos (2009). "A Professor Like Me: The Influence of Instructor Gender on College Achievement." In: *Journal of Human Resources* 44.2, pp. 479–494.

Hoffmann, F. and P. Oreopoulos (2009). "Professor Qualities and Student Achievement." In: *The Review of Economics and Statistics* 91.1, pp. 83–92.

Holmlund, H. and K. Sund (2008). "Is the Gender Gap in School Performance Affected by the Sex of the Teacher?" In: *Labour Economics* 15.1, pp. 37–53.

Jencks, C. and M. Phillips (1998). "The Black-White Test Scope Gap: Why It Persists and What Can Be Done." In: *The Brookings Review*, pp. 24–27.

Kokkelenberg, E. C., M. Dillon and S. M. Christy (2008). "The Effects of Class Size on Student Grades at a Public University." In: *Economics of Education Review* 27.2, pp. 221–233.

Lahelma, E. (2000). "Lack of Male Teachers: A Problem for Students or Teachers?" In: *Pedagogy, Culture and Society* 8.2, pp. 173–186.

Lavy, V. and E. Sand (2015). *On The Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases*. Tech. rep. National Bureau of Economic Research.

Lavy, V. and A. Schlosser (2011). "Mechanisms and Impacts of Gender Peer Effects at School." In: *American Economic Journal: Applied Economics*, pp. 1–33.

Lindo, J. M., N. J. Sanders and P. Oreopoulos (2010). "Ability, Gender, and Performance Standards: Evidence from Academic Probation." In: *American Economic Journal: Applied Economics* 2.2, pp. 95–117.

Marvasti, A. (2007). "Foreign-born teaching assistants and student achievement: an ordered probit analysis." In: *The American Economist* 51.2, pp. 61–71.

Neumark, D. and R. Gardecki (1998). "Women Helping Women? Role Model and Mentoring Effects on Female Ph. D. Students in Economics." In: *Journal of Human Resources*, pp. 220–246.

Nixon, L. A. and M. D. Robinson (1999). "The Educational Attainment of Young Women: Role Model Effects of Female High School Faculty." In: *Demography* 36.2, pp. 185–194.

Oreopoulos, P. and U. Petronijevic (2013). "Making college worth it: A review of the returns to higher education." In: *The Future of Children* 23.1, pp. 41–65.

Rothstein, D. S. (1994). "Do Female Faculty Influence Female Students' Educational and Labor Market Attainments." In: *Indus. & Lab. Rel. Rev.* 48, p. 515.

Ryan, C. L. and K. Bauman (2016). "Educational Attainment in the United States: 2015." In: *March Current Population Report*.

Saft, E. W. and R. C. Pianta (2001). "Teachers' Perceptions of Their Relationships with Students: Effects of Child Age, Gender, and Ethnicity of Teachers and Children." In: *School Psychology Quarterly* 16.2, p. 125.

Schanzenbach, D. W. (2014). "Does Class Size Matter?" In: *Policy Briefs, National Education Policy Center, School of Education, University of Colorado, Boulder*.

Spencer, S. J., C. M. Steele and D. M. Quinn (1999). "Stereotype Threat and Women's Math Performance." In: *Journal of Experimental Social Psychology* 35.1, pp. 4–28.

Turner, S. (2004). "Going to College and Finishing College. Explaining Different Educational Outcomes." In: *College Choices: The Economics of Where to Go, When to Go, and How to Pay for it*, pp. 13–62.

Weston, L. (2014). “OECD: The US Has Fallen Behind Other Countries In College Completion.” In:

Winters, M. A. et al. (2013). “The Effect of Same-gender Teacher Assignment on Student Achievement in the Elementary and Secondary Grades: Evidence from Panel Data.” In: *Economics of Education Review* 34, pp. 69–75.

Main Tables and Figures

Table 1: Descriptive Statistics

	Mean	SD	Obsevations
Panel A. Sample characteristics, student level			19,522
Male	0.528	0.499	
High school Grade Point Average (GPA)	3.641	0.360	
Admitted as transfer	0.201	0.400	
International student	0.039	0.192	
First generation college student	0.388	0.487	
Panel B. Sample characteristics, class level			614
Number of students registered	117.417	83.147	
Professor White	0.713	0.453	
Professor Asian	0.138	0.346	
	White	Asian	Other/ Minority
Panel C. Student outcomes, student-class level			
Dropped class	0.010	0.010	0.019
Observations: 60,642	(0.100)	(0.100)	(0.135)
Numeric course grade (range: 0 to 4)	2.552	2.756	2.348
Observations: 57,718	(1.015)	(0.988)	(1.074)
Passed class	0.840	0.877	0.780
Observations: 59,121	(0.367)	(0.329)	(0.414)
Enroll in an Economics class in future term	0.643	0.605	0.608
Observations: 60,642	(0.479)	(0.489)	(0.488)

	Students			Teaching Assistants		
	Mean	SD	Obs.	Mean	SD	Obs.
Panel D. Student and TA shares by race						
White	0.332	0.471		0.378	0.486	
Asian	0.449	0.497	19,522	0.441	0.597	286
Other/Minority	0.219	0.414		0.182	0.386	

Notes: Panel A corresponds to student-level, Panel B to class-level, and Panel C to student-class level descriptive statistics. In Panel D, reports under "Students" are student-level and reports under "Teaching Assistants" are TA-level. In Panel C, standard deviations are presented in parentheses. The student outcome variable "Numeric course grade" corresponds to the standard numerical American grading system with +/- modifiers (e.g., $A = \frac{12}{3} = 4.0$, $A- = \frac{11}{3} \approx 3.7$, $F = 0$).

Table 2: Tests for Endogenous Sorting – Regression of TA Race on Observables

	Full Sample		Professor Survey Sample	
	(1)	(2)	(3)	(4)
Outcome: Fraction TAs Asian				
Fraction of Students Asian	-0.026	—	0.355	—
	(0.266)	—	(0.402)	—
Fraction of Student Female	-0.351	-1.526	-0.866**	-1.228
	(0.286)	(1.132)	(0.437)	(1.883)
Fraction of Students Admitted as Transfer	0.985*	-1.682	1.072	-3.998
	(0.530)	(1.938)	(0.756)	(3.324)
Average Age of Students	-0.076	0.131	-0.005	0.482
	(0.103)	(0.302)	(0.157)	(0.469)
Fraction of Students International	0.122	2.938	-0.161	1.070
	(0.539)	(2.384)	(0.719)	(3.660)
Fraction of Students First Generation	-0.476	-0.663	-0.982**	-3.447
	(0.325)	(1.298)	(0.469)	(2.402)
Average High School GPA of Students	-0.115	-1.365	0.323	-1.940
	(0.417)	(2.069)	(0.611)	(3.067)
Average Admission Year	0.097	0.028	0.721	0.718
	(0.351)	(0.358)	(0.492)	(0.517)
Fraction of Student Econ Major	-0.109	0.581	-0.161	2.914
	(0.276)	(0.912)	(0.415)	(1.769)
Fraction of Students Double Major	-0.142	0.087	0.167	-1.041
	(0.399)	(1.530)	(0.593)	(2.913)
Average # of Units Up to Class	0.009	-0.006	0.020*	-0.009
	(0.008)	(0.014)	(0.011)	(0.023)
Female Professor	-0.061	-0.149	-0.083	-0.161
	(0.048)	(0.213)	(0.071)	(0.368)
Asian Professor	0.104*	0.013	0.151	-0.012
	(0.056)	(0.290)	(0.099)	(0.673)
Fraction Asian X Fraction Female		2.229		0.575
		(2.168)		(3.545)
Fraction Asian X Fraction Admit as Transfer		5.503		10.345
		(3.772)		(6.406)
Fraction Asian X Average Age		-0.449		-0.932
		(0.558)		(0.867)
Fraction Asian X Fraction International		-4.957		-2.728
		(4.170)		(6.543)
Fraction Asian X Fraction First Generation		0.412		4.849
		(2.506)		(4.539)
Fraction Asian X Average High School GPA		2.450		4.627
		(4.137)		(5.962)
Fraction Asian X Average Admission Year		-0.022		-0.184
		(0.104)		(0.177)
Fraction Asian X Fraction Same Major as Class		-1.401		-5.928*
		(1.803)		(3.355)
Fraction Asian X Fraction Double Major		-0.387		2.387
		(2.968)		(5.335)
Fraction Asian X Average # of Units Up to Class		0.028		0.056
		(0.023)		(0.036)
Fraction Asian X Female Professor		0.193		0.167
		(0.425)		(0.713)
Fraction Asian X Asian Professor		0.195		0.353
		(0.585)		(1.372)
Course & Term FE	Yes	Yes	Yes	Yes
P-value: Joint Significance	0.261	0.709	0.251	0.509
R-squared	0.183	0.190	0.296	0.319
Observations	614	614	334	334

Notes: Each column presents results for a regression where the dependent variable is the fraction of the class's TAs that were Asian. Coefficients for term and course FE are not shown. P-value for joint significance of all individual covariates, conditional on term and course FE, included. The first column includes all baseline characteristics for the class. The second column includes interactions of the fraction of students who were Asian in the class with the baseline characteristics for the class. The next two columns repeat this analyses but for the sample of classes taught by professors who participated in our survey. Standard errors in parentheses. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 3: Main Results

	Standardized grade				Passed class			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full Sample								
Fraction TAs Asian	-0.037***	-0.045***	—	—	-0.013*	-0.014**	—	—
	(0.013)	(0.013)	—	—	(0.007)	(0.007)	—	—
Effect of Similar Race	0.061**	0.076***	0.078***	0.077***	0.017**	0.020**	0.017**	0.030***
$[Asian_i * AsianTA_{ikt}]$	(0.025)	(0.024)	(0.025)	(0.021)	(0.008)	(0.008)	(0.008)	(0.008)
Observations	57718	57718	57718	49177	59121	59121	59121	50329
Panel B: Professor Survey Sample								
Fraction TAs Asian	-0.030	-0.040**	—	—	0.000	-0.002	—	—
	(0.019)	(0.020)	—	—	(0.009)	(0.009)	—	—
Effect of Similar Race	0.066*	0.079**	0.080**	0.086**	0.012	0.014	0.014	0.025**
$[Asian_i * AsianTA_{ikt}]$	(0.034)	(0.033)	(0.034)	(0.034)	(0.011)	(0.010)	(0.011)	(0.012)
Observations	33997	33997	33997	29262	34751	34751	34751	29885
Term FE	X	X			X	X		
Course FE	X	X			X	X		
Class FE			X	X			X	X
Student FE				X				X
Controls:								
Professor		X				X		
Student		X	X			X	X	
Student X Term		X	X	X		X	X	X

Notes: "Effect of Similar Race" is Asian graduate TA composition interacted with an Asian student dummy. Standardized grade has a mean of zero and a standard deviation of one by class. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in Economics, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. All specifications include course-by-race and term-by-race fixed effects. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 4: Audit Study of TA Section and Office Hour Attendance

	Discussion Section			Office Hours			Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome: % students Asian								
Hosted by Asian TA	0.076**	0.085*	0.084*	0.330***	0.134	0.200*	0.081*	0.081**
	(0.034)	(0.044)	(0.045)	(0.120)	(0.124)	(0.103)	(0.045)	(0.038)
Observations		118			43		161	
Mean of outcome		0.576			0.622		0.588	
Controls		X	X		X	X	X	X
Weighted observations			X			X		X

Notes: Each cell reports the coefficient on an indicator for whether the TA for the discussion section or office hour was Asian. The outcome variable is the fraction of attended students who were Asian. Controls include indicators for day of the week, time slot, and auditor. Weights reflect total attendance of the discussion section or office hour. Robust standard errors presented in parenthesis. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 5: Professor Survey Results

		Exams shared w/ TAs?		Multiple choice exams?	
	All	No	Yes	No	Yes
Outcome: Standardized Grade					
Effect of Similar Race	0.080**	-0.003	0.127**	0.199***	0.046
$[Asian_i * AsianTA_{ikt}]$	(0.036)	(0.048)	(0.055)	(0.042)	(0.056)
Observations	33946	9189	19119	9185	24290
Class FE	X	X	X	X	X
Controls:					
Professor					
Student	X	X	X	X	X
Student X Term	X	X	X	X	X

Notes: Standardized grade has a mean of zero and a standard deviation of one by class. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in Economics, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. All specifications include course-by-race and term-by-race fixed effects. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 6: Specifications with Future and Past TA Race

	Standardized Grade						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Future TAs							
$Asian_i \times AsianTA_{ikt}$	0.062*	0.114***	0.086***	0.116***	0.122***	0.078***	0.128***
	(0.034)	(0.028)	(0.030)	(0.028)	(0.027)	(0.024)	(0.025)
$Asian_i \times AsianTA_{i(t+1)}$	-0.006	0.017	0.015	0.010	0.030	0.019	0.020
	(0.030)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Observations	36294	36294	36294	36294	36294	36294	36294
Panel B: Past TAs							
$Asian_i \times AsianTA_{ikt}$	0.116***	0.141***	0.122***	0.139***	0.123***	0.097***	0.122***
	(0.026)	(0.025)	(0.026)	(0.025)	(0.025)	(0.022)	(0.023)
$Asian_i \times AsianTA_{i(t-1)}$	-0.012	-0.014	-0.005	-0.020	-0.005	0.005	0.006
	(0.029)	(0.026)	(0.028)	(0.027)	(0.030)	(0.031)	(0.030)
Observations	37214	37214	37214	37214	37214	37214	37214
Panel C: Sequence Courses							
$Asian_i \times AsianTA_{ikt}$	0.160*	0.156**	0.145*	0.161**	0.069	-0.023	0.097
	(0.081)	(0.067)	(0.074)	(0.064)	(0.100)	(0.118)	(0.106)
$Asian_i \times AsianTA_{prev.course}$	0.105*	0.088*	0.118**	0.081	0.110	0.111	0.104
	(0.057)	(0.047)	(0.050)	(0.050)	(0.089)	(0.088)	(0.095)
Observations	7430	7430	7430	7430	7430	7430	7430
Term FE	X		X			X	
Course FE	X		X			X	
Class FE				X			X
Student FE					X	X	X
Controls:							
Professor		X	X				
Student		X	X	X			
Student X Term		X	X	X			

Notes: In Panels A and B, the sample includes students who enrolled in at least two terms of Economics courses. In Panel C, the sample includes students who enrolled in the second course in of a two-course series for students who enrolled in the entirety of the two-course series. There are a total of nine two-course series that appear in our data: (Introductory/Intermediate) Microeconomics, (Introductory/Intermediate) Macroeconomics, (Introductory/Intermediate) Econometrics, Industrial Organization (A/B), Labor Economics (A/B), World Economic History (A/B), US Economic History (A/B), International Economics (A/B), and Public Economics (A/B). Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 7: Main Results for Enrollment in Another Economics Class and Major in Economics

	Sample of interest					
	Full	Fr., So.	Juniors	Seniors+	Econ Major	Other major
Enroll in Another Class						
Effect of Similar Race	0.008	0.021	0.001	0.004	0.003	0.020
$[Asian_i * AsianTA_{ikt}]$	(0.008)	(0.017)	(0.013)	(0.016)	(0.007)	(0.016)
Observations	60642	22288	19009	19328	32972	27670

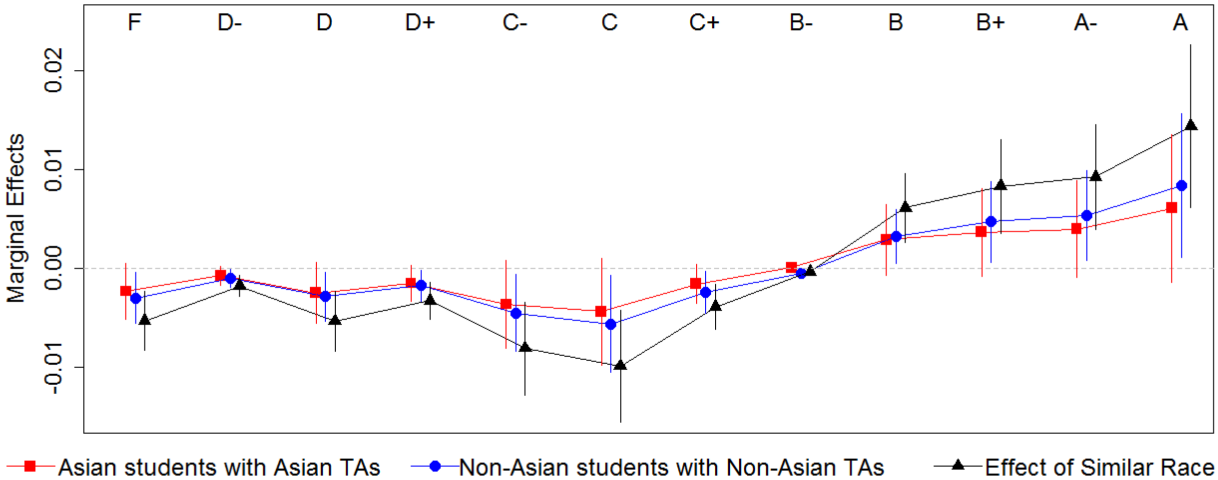
	2+ gen.	First gen.	Exams Withheld	Exams Shared	No MC	Some/All MC
Enroll in Another Class						
Effect of Similar Race	0.003	0.023	0.005	0.032*	0.035**	0.019
$[Asian_i * AsianTA_{ikt}]$	(0.010)	(0.016)	(0.022)	(0.017)	(0.016)	(0.014)
Observations	37232	23410	9674	19994	9765	25323

	Full	Fr., So.	Juniors	Seniors+	Intro Class	Adv. Class
Major in Economics						
Effect of Similar Race	0.007	0.040*	0.004	-0.017	0.014	0.006
$[Asian_i * AsianTA_{ikt}]$	(0.010)	(0.023)	(0.017)	(0.014)	(0.026)	(0.010)
Observations	60642	22288	19009	19328	22366	38276

	2+ gen.	First gen.	Exams Withheld	Exams Shared	No MC	Some/All MC
Major in Economics						
Effect of Similar Race	-0.002	0.028*	0.005	-0.005	-0.012	0.015
$[Asian_i * AsianTA_{ikt}]$	(0.012)	(0.015)	(0.034)	(0.014)	(0.014)	(0.020)
Observations	37232	23410	9674	19994	9765	25323
Class FE	X	X	X	X	X	X
Controls:						
Professor						
Student	X	X	X	X	X	X
Student X Term	X	X	X	X	X	X

Notes: "Effect of Similar Race" is Asian graduate TA composition interacted with an Asian student dummy. "Enroll in Another Class" is an indicator for whether the student enrolled in another class in a future term. "Major in Economics" is an indicator for whether the student "finished" with an Economics major. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in Economics, double major, and class standing (Freshman/Sophomore/Junior/Senior). All specifications include course-by-race and term-by-race fixed effects. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Figure 1: Marginal EFfects of TA Race on Letter Grades



Notes: Each of the three connected lines present marginal effects from an ordered logit regression of having TAs of similar race (or for "Effect of Similar Race", $Asian_i * AsianTA_{ikt}$) on the probability of attaining each possible letter grade. Coefficients reported in Appendix Table A.9. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. Standard errors are clustered by professor. 95% confidence intervals plotted for each estimated marginal effect.

Results Appendix

"Sorting Regressions" from Fairlie et al., 2014:

$$\bar{X}_{ac} = \delta_1 AsianTA_c + \delta_2 I_a + \delta_3 AsianTA_c * I_a + v_{ac} \tag{3}$$

Table A.1: Sorting Regressions – Fairlie et al., 2014 (AER)

	Female	High School GPA	Age	Admit as Transfer	# Prior Units	Double Major	Inter-national	Class Major
Full Sample	0.014	-0.018	-0.024	0.007	-1.537	-0.014	-0.000	0.003
	(0.018)	(0.014)	(0.061)	(0.015)	(1.924)	(0.018)	(0.012)	(0.017)
Professor Survey Subsample	0.013	-0.001	-0.065	0.008	-1.759	-0.008	0.009	-0.009
	(0.023)	(0.016)	(0.087)	(0.022)	(2.934)	(0.027)	(0.018)	(0.022)
Classes with one TA	0.005	-0.019	-0.060	-0.007	-1.390	-0.011	-0.003	0.002
	(0.024)	(0.019)	(0.071)	(0.020)	(2.424)	(0.024)	(0.016)	(0.021)
Class FE	X	X	X	X	X	X	X	X

Notes: Each cell displays results from a regression of the race-specific average student outcomes in a classroom on an indicator for whether the average is associated with Asian students, the fraction of the TAs assigned to the class who are Asian, the interaction between these two variables, and class fixed effects. This table reports the coefficient on the interaction term, which can be interpreted as the extent to which Asian students sort into classes assigned Asian TAs. Outcomes for each regression vary across columns. Rows are defined by the subsample of students we consider. Students and TAs are classified as Asian if their primary race is recorded as Chinese, Japanese, Korean, Filippino, South-East Asian, Vietnamese, Thai, or "Other Asian". Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.2: Regression of TA Race on Observables - Sample of students who completed course

	Full Sample		Prof. Survey Sample	
	(1)	(2)	(3)	(4)
Outcome: Fraction TAs Asian				
Fraction of Students Asian	-0.001	—	0.301	—
	(0.263)	—	(0.395)	—
Fraction of Student Female	-0.299	-1.457	-0.598	-1.199
	(0.274)	(1.061)	(0.418)	(1.798)
Fraction of Students Admitted as Transfer	1.152**	-2.221	1.266*	-4.090
	(0.517)	(1.855)	(0.735)	(3.186)
Average Age of Students	-0.113	0.174	-0.060	0.346
	(0.101)	(0.296)	(0.154)	(0.458)
Fraction of Students International	0.094	2.492	-0.201	1.320
	(0.522)	(2.185)	(0.687)	(3.373)
Fraction of Students First Generation	-0.529*	-0.606	-0.962**	-3.264
	(0.318)	(1.263)	(0.458)	(2.369)
Average High School GPA of Students	0.045	-1.745	0.296	-2.121
	(0.414)	(2.031)	(0.597)	(3.096)
Average Admission Year	0.049	-0.000	0.538	0.570
	(0.347)	(0.355)	(0.483)	(0.510)
Fraction of Student Econ Major	-0.134	0.737	-0.152	3.382**
	(0.272)	(0.869)	(0.411)	(1.621)
Fraction of Students Double Major	-0.171	0.458	0.200	-1.190
	(0.393)	(1.479)	(0.584)	(2.818)
Average # of Units Up to Class	0.008	-0.012	0.017	-0.012
	(0.008)	(0.013)	(0.010)	(0.022)
Female Professor	-0.060	-0.133	-0.083	-0.082
	(0.048)	(0.211)	(0.071)	(0.371)
Asian Professor	0.107*	0.003	0.151	-0.041
	(0.056)	(0.284)	(0.099)	(0.670)
Fraction Asian X Fraction Female		2.204		0.944
		(2.041)		(3.403)
Fraction Asian X Fraction Admit as Transfer		6.817*		10.947*
		(3.594)		(6.152)
Fraction Asian X Average Age		-0.572		-0.774
		(0.543)		(0.850)
Fraction Asian X Fraction International		-4.359		-3.389
		(3.771)		(6.012)
Fraction Asian X Fraction First Generation		0.153		4.594
		(2.434)		(4.476)
Fraction Asian X Average High School GPA		3.604		4.974
		(4.044)		(5.986)
Fraction Asian X Average Admission Year		-0.013		-0.169
		(0.102)		(0.176)
Fraction Asian X Fraction Same Major as Class		-1.821		-6.873**
		(1.718)		(3.081)
Fraction Asian X Fraction Double Major		-1.179		2.771
		(2.887)		(5.187)
Fraction Asian X Average # of Units Up to Class		0.037*		0.058*
		(0.022)		(0.034)
Fraction Asian X Female Professor		0.163		0.013
		(0.420)		(0.716)
Fraction Asian X Asian Professor		0.213		0.410
		(0.570)		(1.361)
Course & Term FE	Yes	Yes	Yes	Yes
P-value: Joint Significance	0.219	0.606	0.315	0.422
R-squared	0.184	0.193	0.294	0.323
Observations	614	614	334	334

Notes: Each column presents results for a regression where the dependent variable is the fraction of the class's TAs that were Asian. Coefficients for term and course FE are not shown. P-value for joint significance of all individual covariates, conditional on term and course FE, included. The first column includes all baseline characteristics for the class. The second column includes interactions of the fraction of students who were Asian in the class with the baseline characteristics for the class. The next two columns repeat this analyses but for the sample of classes taught by professors who participated in our survey. Standard errors in parentheses. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.3: Tests for Sorting – Regression of TA Race on Observables – Student-class data

	Full Sample			Prof. Survey Sample		
	All students	Asian	Non-Asian	All	Asian	Non-Asian
Outcome: Fraction TAs Asian						
Asian Student	0.001	—	—	0.002	—	—
	(0.004)	—	—	(0.004)	—	—
Female Student	0.001	0.003	-0.003	-0.001	0.005	-0.006
	(0.003)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Admit as Transfer	0.016**	0.023*	0.013	0.011	0.023	0.004
	(0.008)	(0.012)	(0.011)	(0.010)	(0.015)	(0.013)
Age	-0.001	-0.002	0.001	0.001	-0.002	0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
International Student	-0.011	-0.009	-0.014	-0.014	-0.009	-0.030
	(0.010)	(0.010)	(0.018)	(0.010)	(0.011)	(0.021)
First Generation	-0.003	-0.002	-0.004	-0.005	-0.004	-0.006
	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)
High School GPA	0.004	-0.006	0.013**	0.003	-0.001	0.008
	(0.005)	(0.008)	(0.006)	(0.007)	(0.009)	(0.008)
Admission Year	-0.005	-0.004	-0.007	-0.001	-0.003	-0.001
	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Same Major as Class	0.001	0.002	-0.003	0.001	-0.001	-0.001
	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
Double Major	0.001	0.001	0.001	0.007	0.005	0.010
	(0.007)	(0.010)	(0.009)	(0.009)	(0.011)	(0.011)
Freshman	0.002	-0.010	0.012	-0.001	-0.015	0.012
	(0.010)	(0.014)	(0.012)	(0.013)	(0.016)	(0.015)
Sophomore	0.008	0.005	0.009	0.002	-0.001	0.007
	(0.007)	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)
Junior	0.001	0.001	0.003	-0.002	-0.010	0.007
	(0.006)	(0.008)	(0.007)	(0.008)	(0.010)	(0.008)
Female Professor	-0.056	-0.057	-0.056	-0.031	-0.047	-0.016
	(0.042)	(0.041)	(0.044)	(0.059)	(0.058)	(0.061)
Asian Professor	0.054	0.046	0.061	0.093	0.081	0.107
	(0.047)	(0.047)	(0.049)	(0.071)	(0.073)	(0.071)
Course & Term FE	Yes	Yes	Yes	Yes	Yes	Yes
P-value: Joint Significance	0.706	0.421	0.389	0.891	0.540	0.592
R-squared	0.220	0.235	0.209	0.306	0.311	0.306
Observations	60642	29391	31251	35023	17448	17575

Notes: Each specification presents results for a regression where the dependent variable is the fraction of the student's TAs in the class that were Asian. Coefficients for term and course FE are not shown. P-value for joint significance of all individual covariates, conditional on term and course FE, included. The first column is our full sample. The next two columns consider Asian and non-Asian student subsamples. The final three columns pertain to the sample of classes taught by professors who participated in our survey. Standard errors in parentheses, clustered by class. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.4: Main Results by White vs. Non-White

	Standardized grade				Passed class			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full Sample								
Fraction TAs White	-0.031***	-0.029***	—	—	-0.003	-0.002	—	—
	(0.010)	(0.011)	—	—	(0.007)	(0.007)	—	—
Effect of Similar Race	0.111***	0.104***	0.107***	0.076***	0.022***	0.021**	0.020**	0.022***
$[White_i * WhiteTA_{ikt}]$	(0.026)	(0.024)	(0.024)	(0.021)	(0.008)	(0.008)	(0.008)	(0.008)
Observations	57718	57718	57718	49177	59121	59121	59121	50329
Panel B: Professor Survey Sample								
Fraction TAs White	-0.041***	-0.038**	—	—	-0.010	-0.008	—	—
	(0.014)	(0.015)	—	—	(0.009)	(0.009)	—	—
Effect of Similar Race	0.119***	0.109***	0.114***	0.072**	0.024**	0.021**	0.023**	0.026**
$[White_i * WhiteTA_{ikt}]$	(0.035)	(0.032)	(0.033)	(0.032)	(0.010)	(0.010)	(0.010)	(0.011)
Observations	33997	33997	33997	29262	34751	34751	34751	29885
Term FE	X	X			X	X		
Course FE	X	X			X	X		
Class FE			X	X			X	X
Student FE				X				X
Controls:								
Professor		X				X		
Student		X	X			X	X	
Student X Term		X	X	X		X	X	X

Notes: "Effect of Similar Race" is White graduate TA composition interacted with an White student dummy. Standardized grade has a mean of zero and a standard deviation of one by class. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in Economics, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. All specifications include course-by-race and term-by-race fixed effects. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.5: Estimated Role of TA Race for Student Outcomes - Group by Group Regressions

	Outcome: Standardized Grade				Outcome: Passed Class			
	Racial Composition of TAs (Comparison Group: Own Race TA)				*Racial Composition of TAs (Comparison Group: Own Race TA)*			
	White	Chinese	*Other Asian*	*Hispanic (Other)*	*White*	Chinese	*Other Asian*	*Hispanic (Other)*
White	—	-0.078***	-0.081***	-0.041	—	-0.016*	-0.022***	-0.018
	—	(0.017)	(0.022)	(0.032)	—	(0.008)	(0.008)	(0.012)
Chinese	-0.073***	—	-0.008	-0.027	-0.014	—	0.004	-0.022
	(0.024)	—	(0.022)	(0.040)	(0.009)	—	(0.009)	(0.017)
Other Asian	-0.008	-0.034	—	0.064	0.003	-0.012	—	0.017
	(0.032)	(0.030)	—	(0.049)	(0.012)	(0.013)	—	(0.024)
Hispanic (Other)	0.009	0.043	0.003	—	-0.012	-0.022	-0.009	—
	(0.052)	(0.050)	(0.059)	—	(0.021)	(0.019)	(0.022)	—

Notes: This table displays results from regressions that are run separately for each student race. Each cell reports the coefficient for TA racial composition. Standardized grade has a mean of zero and a standard deviation of one. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. Course and term fixed effects included. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.6: Full Model with Multiple Same Race Interactions

	Standardized Grade						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Chinese_i \times ChineseTA_{ikt}$	0.051	0.063**	0.065**	0.067**	0.072***	0.076***	0.084***
	(0.038)	(0.031)	(0.030)	(0.031)	(0.025)	(0.025)	(0.027)
$OtherAsian_i \times OtherAsianTA_{ikt}$	0.032	0.056	0.053	0.053	0.016	0.021	0.017
	(0.038)	(0.035)	(0.035)	(0.035)	(0.031)	(0.030)	(0.029)
$White_i \times WhiteTA_{ikt}$	0.094***	0.114***	0.105***	0.107***	0.089***	0.086***	0.086***
	(0.029)	(0.026)	(0.026)	(0.027)	(0.022)	(0.020)	(0.021)
Observations	57718	57718	57718	57718	49177	49177	49177
Term FE	X		X			X	
Course FE	X		X			X	
Class FE				X			X
Student FE					X	X	X
Controls:							
Professor		X	X				
Student		X	X	X			
Student X Term		X	X	X			

Notes: Each cell reports the coefficient on the interaction between a student race identifier and the fraction of TAs who were of similar race. For each regression, covariates for students/TAs of "Other" race are omitted. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, and class standing (Freshman/Sophomore/Junior/Senior). Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.7: Robustness Checks – Additional Specifications with Standarfized Grade

	Standardized Grade						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full Sample							
Fraction TAs Asian	-0.037**	-0.079***	-0.045***	—	-0.132***	-0.040**	—
	(0.014)	(0.015)	(0.014)	—	(0.017)	(0.016)	—
Effect of Similar Race	0.061**	0.098***	0.076***	0.096***	0.113***	0.074***	0.116***
$[Asian_i * AsianTA_{ikt}]$	(0.026)	(0.023)	(0.024)	(0.024)	(0.021)	(0.020)	(0.020)
Observations	57718	57718	57718	57718	49177	49177	49177
Panel B: Professor Survey Sample							
Fraction TAs Asian	-0.030	-0.072***	-0.041*	—	-0.159***	-0.063**	—
	(0.023)	(0.020)	(0.023)	—	(0.022)	(0.023)	—
Effect of Similar Race	0.065*	0.083***	0.077**	0.079**	0.120***	0.081***	0.118***
$[Asian_i * AsianTA_{ikt}]$	(0.038)	(0.031)	(0.036)	(0.031)	(0.027)	(0.026)	(0.025)
Observations	33399	33399	33399	33399	28683	28683	28683
Panel C: Single TA Class							
Fraction TAs Asian	-0.051***	-0.070***	-0.052***	—	-0.077**	-0.037*	—
	(0.018)	(0.021)	(0.017)	—	(0.030)	(0.022)	—
Effect of Similar Race	0.090***	0.106***	0.096***	0.108***	0.109***	0.100***	0.113***
$[Asian_i * AsianTA_{ikt}]$	(0.031)	(0.034)	(0.030)	(0.035)	(0.033)	(0.030)	(0.034)
Observations	17500	17500	17500	17500	16841	16841	16841
Term FE	X		X			X	
Course FE	X		X			X	
Class FE				X			X
Student FE					X	X	X
Controls:							
Professor		X	X				
Student		X	X	X			
Student X Term		X	X	X			

Notes: "Effect of Similar Race" is Asian graduate TA composition interacted with an Asian student dummy. Standardized grade has a mean of zero and a standard deviation of one by class. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.8: Tests for Attrition – Dummy for Dropped Class Regressed on Observables Interacted w/ TA Race

	Dropped Class					
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction TAs Asian interacted w/						
Student Asian	0.002	0.002	0.002	0.003	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Student Female	-0.001	-0.001	-0.000	-0.001	-0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Admitted as Transfer	-0.008*	-0.009*	-0.008	-0.016***	-0.018***	-0.020***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Student International	0.004	0.003	0.003	0.000	-0.001	-0.002
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)
Student First Generation	-0.002	-0.002	-0.002	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
High School GPA	-0.001	-0.001	-0.000	-0.004	-0.004	-0.003
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Econ Major	-0.001	-0.002	0.001	0.000	0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Student Double Major	0.003	0.003	0.002	0.005	0.005	0.004
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Student Age	0.002***	0.002**	0.002***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
# of Units Up to Class	-0.000	-0.000	-0.000	-0.000	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female Professor	0.001	0.003		0.003	0.002	
	(0.003)	(0.003)		(0.003)	(0.002)	
Asian Professor	0.001	0.002		0.001	0.002	
	(0.004)	(0.003)		(0.005)	(0.003)	
Observations	60642	60642	60642	51653	51653	51653
Term FE		X			X	
Course FE		X			X	
Class FE			X			X
Student FE				X	X	X

Notes: Each column presents results from a single OLS regression where the outcome is an indicator for whether the student dropped the class at any point during the term. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively. We find positive and statistically insignificant responses in class drop rates for students who are matched to TAs of similar race.

Table A.9: Alternative Specifications using Letter Grades

Panel A: Coefficients

	Ordered Logit	Ordered Probit
Asian Students		
w/ Asian TAs	0.067	0.034
	(0.042)	(0.024)
Non-Asian Students		
w/ Non-Asian TAs	0.087**	0.050**
	(0.038)	(0.022)
Effect of Similar Race	0.153***	0.083***
$[Asian_i * AsianTA_{ikt}]$	(0.045)	(0.026)
N	57,718	57,718
log L	-131532.85	-131525.46
Pseudo R^2	0.0326	0.0327

Panel B: Marginal Effects

	F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A
Ordered Logit												
Asian Students												
w/ Asian TAs	-0.002	-0.001	-0.002	-0.002	-0.004	-0.004	-0.002	0.001	0.003	0.004	0.004	0.006
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)
Non-Asian Students												
w/ Non-Asian TAs	−0.003**	−0.001**	−0.003**	−0.002**	−0.005**	−0.006**	−0.002**	0.001**	0.003**	0.005**	0.005**	0.008**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)
Effect of Similar Race	−0.005***	−0.002***	−0.005***	−0.003***	−0.008***	−0.010***	−0.004***	0.001***	0.006***	0.008***	0.009***	0.014***
$[Asian_i * AsianTA_{ikt}]$	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)
Ordered Probit												
Asian Students												
w/ Asian TAs	-0.003	-0.001	-0.002	-0.001	-0.003	-0.003	-0.001	0.001	0.002	0.003	0.003	0.006
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)
Non-Asian Students												
w/ Non-Asian TAs	−0.004**	−0.001**	−0.003**	−0.002**	−0.004**	−0.004**	−0.002**	0.001**	0.002**	0.004**	0.005**	0.010**
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)
Effect of Similar Race	−0.006***	−0.002***	−0.005***	−0.003***	−0.007***	−0.007***	−0.003***	0.001***	0.004***	0.006***	0.008***	0.015***
$[Asian_i * AsianTA_{ikt}]$	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.005)

Notes: Panel A shows the estimated coefficients from ordered logit and ordered probit models. Panel B displays the marginal effects for each possible letter grade, evaluated at the controls' means. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.10: Robustness of Professor Survey Results

	No Multiple Choice Exams				Some/All Multiple Choice			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome: Standardized Grade								
Fraction TAs Asian	-0.132***	-0.139***	—	—	-0.002	-0.026	—	—
	(0.024)	(0.045)	—	—	(0.030)	(0.033)	—	—
Effect of Similar Race	0.196***	0.201***	0.199***	0.217***	0.046	0.035	0.046	0.040
$[Asian_i * AsianTA_{ikt}]$	(0.041)	(0.065)	(0.042)	(0.070)	(0.055)	(0.044)	(0.056)	(0.048)
Observations	9185	8883	9185	8883	24290	19861	24290	19861
	Exams Withheld from TAs				Exams Shared with TAs			
Fraction TAs Asian	0.000	-0.105	—	—	-0.045	-0.039	—	—
	(0.032)	(0.095)	—	—	(0.035)	(0.045)	—	—
Effect of Similar Race	-0.000	0.075	-0.003	0.056	0.126**	0.155*	0.127**	0.172**
$[Asian_i * AsianTA_{ikt}]$	(0.046)	(0.096)	(0.048)	(0.093)	(0.055)	(0.081)	(0.055)	(0.081)
Observations	9189	8296	9189	8296	19119	15503	19119	15503
Term FE	X	X			X	X		
Course FE	X	X			X	X		
Class FE			X	X			X	X
Student FE		X		X		X		X
Controls:								
Professor	X	X			X	X		
Student	X		X		X		X	
Student X Term	X	X	X	X	X	X	X	X

Notes: "Effect of Similar Race" is Asian graduate TA composition interacted with an Asian student dummy. Standardized grade has a mean of zero and a standard deviation of one by class. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, class standing (Freshman/Sophomore/Junior/Senior), and professor gender and race. All "Effect of Similar Race" specifications include course-by-race and term-by-race fixed effects. Standard errors in parentheses, clustered by professor. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A.11: Additional Professor Survey Results

	Professor Survey Sample					
	All	(0)	(1)	(2)	(3)	(4)
Outcome: Standardized Grade						
Effect of Similar Race	0.080**	0.066	-0.067	0.167	0.179	0.235***
$[Asian_i * AsianTA_{ikt}]$	(0.036)	(0.039)	(0.057)	(0.097)	(0.128)	(0.046)
Multiple Choice Exams	—	Yes & No	Yes	Yes	No	No
Share Exams with TAs	—	Yes & No	No	Yes	No	Yes
Class FE	X	X	X	X	X	X
Course X Race FE	X	X	X	X	X	X
Term X Race FE	X	X	X	X	X	X
Controls:						
Professor						
Student	X	X	X	X	X	X
Student X Term	X	X	X	X	X	X
Observations	33997	27837	7915	13643	897	5382

Notes: Each cell reports the coefficient on the interaction between a student identifier for Asian and fraction of TAs Asian. The first column reports estimates for the subsample of classes where professors completed the survey. Column (0) considers the subsample of classes where professors answered both questions of interest. The remaining columns consider further survey subsamples. Controls include age when class began, high school GPA, and admission year, as well as indicators for student gender, international vs. domestic, whether parents attended college, admittance (transfer vs. freshman), whether the student is majoring in the subject of the course, double major, and class standing (Freshman/Sophomore/Junior/Senior). Standard errors in parentheses, clustered by professor for Panels A and B, and clustered by TA for Panel C. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.