I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10459

Stereotype Threat, Role Models, and Demographic Mismatch in an Elite Professional School Setting

Christopher Birdsall
Seth Gershenson
Raymond Zuniga

DECEMBER 2016

IZA – Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10459

Stereotype Threat, Role Models, and Demographic Mismatch in an Elite Professional School Setting

Christopher Birdsall
Boise State University

Seth Gershenson
American University and IZA

Raymond Zuniga
American University

DECEMBER 2016

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Stereotype Threat, Role Models, and Demographic Mismatch in an Elite Professional School Setting*

Ten years of administrative data from a diverse, private, top-100 law school are used to examine the ways in which female and nonwhite students benefit from exposure to demographically similar faculty in first-year required law courses. Arguably causal impacts of exposure to same-sex and same-race instructors on course-specific outcomes such as course grades are identified using a two-way (student and classroom) fixed effects strategy. Impacts of faculty representation on long-run, student-specific outcomes such as graduation are identified using an instrumental variables (IV) strategy that exploits transitory variation in the demographic makeup of the faculty. Having an other-sex instructor reduces the likelihood of receiving a good grade (A or A-) by one percentage point (3%) and having an other-race instructor reduces the likelihood of receiving a good grade by three percentage points (10%). The effects of student-instructor demographic mismatch are particularly salient for nonwhite female students. The IV estimates suggest that the share of first-year courses taught by nonwhite instructors increases the probabilities that nonwhite students persist into the second year and graduate on time. These results provide novel evidence of the pervasiveness of role-model effects in elite settings and of the graduate-school education production function.

JEL Classification: I23, J15, J44

Keywords: demographic mismatch, law school, gender, race

Corresponding author:
Seth Gershenson
School of Public Affairs
American University
4400 Massachusetts Avenue, NW
Washington, DC 20016-8070
USA
E-mail: gershens@american.edu

* Stephen B. Holt, Billie Jo Kaufman, Michal Kurlaender, Nicholas Papageorge, and participants at the 2016 APPAM Fall Conference and 2016 Access Group Legal Education Research Symposium provided many helpful comments. The authors are thankful for financial support from the Association for Institutional Research (AIR) Research Grant Program. Opinions reflect those of the authors and not necessarily those of the granting agency. Kimberly Trocha provided excellent research assistance.

1 Introduction

An emerging literature in the economics of education documents wide-ranging impacts of student-teacher demographic match on both students and teachers. In K-12 classrooms, assignment to an other-race or other-sex teacher has been shown to harm student achievement (Dee, 2004, 2007).[1] Similarly, racial mismatch lowers teachers' perceptions of student behavior (Dee, 2005) and their expectations for students' educational attainment (Gershenson, Holt and Papageorge, 2016). The impact of faculty representation has also been studied in the post-secondary context, particularly among first-year undergraduates (Bettinger and Long, 2005; Carrell, Page and West, 2010; Hoffmann and Oreopoulos, 2009; Fairlie, Hoffmann and Oreopoulos, 2014). These studies typically find modest effects of having a same-sex or same-race instructor on course grades, the likelihood of dropping a class, and choice of major. Lusher, Campbell and Carrell (2015) show similar effects of having a same-race teaching assistant (recitation section leader) on course grades and office-hour and course attendance.

While the precise mechanisms through which student-instructor demographic mismatch affects students' educational outcomes are not known, it is generally thought that role model effects, stereotype threat, and information provision play prominent roles in this phenomenon. Moreover, it is often, either implicitly or explicitly, assumed that relatively young, inexperienced, socio-economically disadvantaged, and information-poor students are particularly susceptible to the deleterious effects of student-instructor demographic mismatch. In the current study, we show that the harms associated with student-instructor demographic mismatch are just as pronounced in an elite, professional school setting as they are in K-12, community college, and first-year undergraduate classrooms. These findings provide novel evidence that mismatch effects are not limited to inexperienced, disadvantaged, or otherwise vulnerable populations.[2] Rather, student-instructor demographic mismatch continues to harm the academic performance of even elite law school students, whom we might falsely deem impervious to such threats, given that they are college graduates who successfully navigated the law school application process. This suggests that student-instructor mismatch might affect student outcomes through channels over and above those commonly considered.

Specifically, we use rich administrative data from a top-100 law school and an array of arguably causal fixed-effects and instrumental-variable identification strategies to show that having a demographically mismatched first-year law instructor significantly reduces the

[1]Mismatch is not universally harmful, however, as Antecol, Eren and Ozbeklik (2015) find that less-prepared female math teachers reduce female students' achievement, but have no such effect on male students.

[2]There is a litany of qualitative and anecdotal evidence of such demographic biases in legal education (Banks, 1988; Darling-Hammong and Holmquist, 2015; Guinier et al., 1994), but to our knowledge there is no credibly identified, quantitative evidence on the impact of law student-instructor demographic match on student outcomes.

probability of receiving a "good grade" (A/A-) in the course. Other-race effects tend to be larger in magnitude than other-sex effects, particularly among nonwhite female students, though both are statistically and economically significant. There are cumulative effects of exposure to demographically mismatched faculty as well, both of first-semester instructors on second-semester course grades in two-course sequences, and of the share of mismatch experienced in the first-year required courses on longer-run outcomes such as persistence and graduation. Classroom environments such as class size and class composition moderate the impact of student-instructor demographic mismatch in ways that hint at the mechanisms through which such effects operate. That we find such effects in an elite professional school setting suggests that the phenomena of stereotype threat and role-model effects are not solely attributable to a lack of information, confidence, or experience. Rather, these are broader, societal phenomena that permeate beyond relatively vulnerable populations of schoolchildren and community college students and have implications for all social interactions, not just those in which there is a power dynamic (e.g., doctor-patient). Our findings suggest the need to extend our understanding to stereotype threat, role models, and mismatch effects to include other settings and contexts where such effects have been heretofore presumed to play relatively small roles.

A second contribution of the current study is to enhance our understanding of the production of graduate degrees. Indeed, remarkably little is known about the nature of the law-school education production function, or that for graduate school more generally.[3] This is troubling, as graduate students comprise a nontrivial segment of the U.S. post-secondary student population: about 15 percent of post-secondary students are graduate students and about 40 percent of outstanding student-loan debt was accumulated to finance graduate degrees (Delisle, 2014). Graduate degrees themselves facilitate entrance into many high-status and high-paying professions central to the modern economy. The legal profession is one prominent example: nearly all states require that lawyers hold a Juris Doctor (JD) from an American Bar Association (ABA) accredited law school, lawyers constitute about 1% of the U.S. labor force, and law firm revenues constitute about 1% of U.S. GDP (Azmat and Ferrer, Forthcoming). The current study provides evidence on some of the educational inputs and environments that affect law school students' achievement, skill development, choice of specialization, and persistence.

The current study also sheds light on the role that institutions play in perpetuating demographic wage, skill, and partnership gaps in the legal profession. For example, female

[3] Exceptions include recent natural experiments involving first-year law students at Stanford who were randomly assigned to small classes (Ho and Kelman, 2014) and at Minnesota who were randomly assigned to receive individualized feedback (Schwarcz and Farganis, Forthcoming).

lawyers earn lower salaries and are less likely to be promoted to partner than their male counterparts, even after conditioning on basic employee and firm characteristics (Azmat and Ferrer, Forthcoming; Dinovitzer, Reichman and Sterling, 2009; Wood, Corcoran and Courant, 1993).[4] Azmat and Ferrer (Forthcoming) show that performance gaps explain much of the previously unexplained sex gap in lawyers' earnings, though the exact sources of gaps in performance and specialization among practicing lawyers remain unclear. Law school environments and mentoring practices might contribute to this divergence in post-law school productivity, even when male and female students enter law school with similar skills (Bertrand, 2011; Ho and Kelman, 2014). We test this hypothesis by examining whether the demographic match between law students and instructors affects student outcomes. Doing so will inform law-school policy and practice by identifying the malleable factors that influence the success of underrepresented graduate school students and our understanding of the importance that faculty play in the production of graduate education more generally. Indeed, law schools are representative of a broad class of professional graduate schools and programs from which professional service providers are recruited directly into the labor market (e.g., business, engineering) (Oyer and Schaefer, 2015).

Documenting the impact of having an other-race instructor in the law school context is also important due to the social consequences of demographic gaps in the receipt of law degrees and in the career paths of law school graduates (Holder Jr, 2001). For example, the under-representation of racial and ethnic minorities in the U.S. judiciary likely contributes to documented demographic disparities in sentencing (Mustard, 2001). Indeed, implicit association tests (IATs) show that white judges often hold implicit (unconscious) biases against non-white defendants (Rachlinski et al., 2008). In the field, emotional shocks associated with the outcomes of football games have been shown to increase the sentences assigned by judges, particularly for black defendants (Eren and Mocan, 2016). And regarding the demographic pay gaps discussed above, a lack of representation among law school faculty and/or how law school faculty interact with and mentor women and students of color can cause sorting into specializations and other behavioral responses that affect prestige, pay, and upward mobility. Ultimately, biases against females and people of color can produce self-fulfilling prophecies in which members of stereotyped groups ultimately conform to what were initially incorrect beliefs (Papageorge, Gershenson and Kang, 2016; Steele, 1997; Loury, 2009). Institutional factors such as the faculty composition of law schools can therefore perpetuate the under-representation of certain demographic groups in the legal profession (Wilkins and Gulati, 1996).

[4]This is consistent with "glass ceilings" and pay gaps in top management positions (Bertrand and Hallock, 2001), as well as in the labor force more generally (Altonji and Blank, 1999).

The paper proceeds as follows: Section 2 describes the administrative data and institutional details. Section 3 introduces the two-way fixed effects identification strategy. Section 4 presents the main results. Section 5 concludes.

2 Data & Institutional Details

This section describes the administrative data analyzed in the current study. Section 2.1 describes the institutional context and the formation of the analytic sample. Section 2.2 summarizes the analytic sample.

2.1 Administrative Data

All analyses use longitudinal administrative data from a private, top-100 law school (LS) located in a major urban center. The LS enrolls approximately 1,000 students, on average, and employs approximately 200 full- and part-time faculty. It is one of the most demographically and geographically diverse top-ranked law schools, as its student body is majority female, almost 40% nonwhite, and includes students from almost every state in the U.S. as well as from several foreign countries.

The most recent U.S. News rankings rank the LS in the same range as the University of Oregon, University of Pittsburgh, Villanova, University of Denver, and Northeastern law schools.[5] Demographically, LS ranks in the top 50 ABA-approved law schools for racial/ethnic minority JD-student enrollment. Its peer institutions in this category include the University of California-Irvine, the City University of New York, and Rutgers University. Similarly, LS ranks in the top 20 ABA-approved law schools in terms of female JD enrollment. Institutions with similar female enrollments include Boston University, University of California-Davis, and the City University of New York law schools.[6] Thus, while LS is one of the more demographically diverse law schools in the U.S., it is not an outlier and is comparable to other highly-ranked, national law schools in this regard. It also has a relatively diverse faculty, as described below.

The analytic sample is restricted to students' first-year required courses, as entering students take the same set of courses during their first two semesters of law school and these courses are typically scheduled by LS advisors and administrators. The administrative data include detailed information on course-specific outcomes such as grades, dropout behavior,

[5] `http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-law-schools/law-rankings/page+4`.

[6] Rankings calculated as average % enrollment from 2009 to 2013 using data obtained from the American Bar Association: `http://www.americanbar.org/groups/legal_education/resources/statistics.html`

and taking a subsequent course in the same concentration, as well as student-level outcomes such as persistence, graduation, and engagement with the LS's Law Journals, for every student who entered the JD program between fall 2000 and fall 2011. Additionally, we observe student demographic characteristics, such as sex, age, and race/ethnicity, as well as LSAT scores, undergraduate GPA, and home zip code. We use home zip codes to construct measures of distance from LS and to collect the median income and fraction of adults who have a college degree in each zip code from the 2000 and 2010 U.S. censuses, which proxy for students' socioeconomic status. Administrative data on instructors include rank (e.g., tenure line, tenured, adjunct) and years at LS. Demographic information (i.e., race/ethnicity and sex) and rank of faculty members' JD-granting institutions were determined by reviewing public resumes, curriculum vitae, and websites.[7]

2.2 Sample and Summary Statistics

Our aim is to estimate the impact of student-instructor demographic match in first-year required courses. The primary unit of analysis is therefore the student-course level. There are eight required courses in the first year, which cover subjects such as litigation, constitutional law, criminal law, and property law. The main analytic sample includes 37,042 student-course observations from more than 1,000 unique course sections. Panel A of Table 1 summarizes the student-course data, separately by students' race and sex. On average, white students have higher first-year course grades than nonwhite students. There is no appreciable sex-gap in first-year course grades. Dropping first-year required courses is exceedingly rare, likely because they are required and students are generally forbidden from switching sections. White students and nonwhite students have near-equal likelihoods of having an other-sex instructor, while females are more likely than males to have an other-sex instructor. Nonwhite students are much more likely to have an other-race instructor than are white students, which reflects the fact that the majority of instructors are white.

Panel B of Table 1 reports descriptive statistics at the student level. The average age of first-year JD students is about 25 for all demographic groups. While females form a majority of both white and nonwhite students, the representation of females is greater among nonwhite students than among white students. White and male students tend to have higher LSAT scores than nonwhite and female students. Among non-white students, 21% are black, 37% are Latino, and 35% are Asian. Graduation rates are similar across demographic groups.

Finally, Panel C of Table 1 reports descriptive statistics at the instructor level, for instructors who taught at least one first-year required course between 2000 and 2012. On

[7]The rank of instructors' JD programs comes from the usual US News Rankings.

average, white instructors have more experience at LS than nonwhite instructors, and male instructors have more experience than female instructors. About 46% percent of white instructors are female, while 60% of non-white instructors are female. Almost half of nonwhite faculty are black, 23% are Latino, and 29% are Asian. More than 90% of instructors have a JD, though some have other advanced degrees.[8] The average instructor attended a JD program ranked in the top 40 by US News. White and male instructors attended slightly higher ranked programs, on average, than did nonwhite and female instructors, respectively.

Table 2 reports descriptive statistics at the classroom (i.e., course-section) level. There are 1,132 unique first-year required course offerings in the analytic sample. The average class contained about 43 students, 59% of whom were female. The majority (86%) of courses were taught by white faculty, while 8% were taught by black instructors, 4% by Latino instructors, and 2% by Asian instructors. Table 2 also reports the frequency of the 12 courses that were at one time required in the first year between 2000 and 2012. Some courses appear less often either because some courses were merged into a single course or they ceased to be required between 2000 and 2012.

3 Identification Strategy

This section describes the main identification strategy used to estimate the causal effects of student-instructor demographic match on course-specific outcomes. Section 3.1 introduces the preferred two-way fixed effects specification. Section 3.2 discusses the key identifying assumptions and presents a test of the "endogenous sorting" threat to identification. Finally, Section 3.3 describes a three-way fixed effects specification used to identify the effect of mismatch in the first course of two-course sequences on performance in the second course.

3.1 Baseline Model

Our primary interest is in how student-instructor demographic match affects outcomes (y) at the student-course level. Specifically, we are interested in δ in the linear regression model:

$$y_{ijcst} = \beta_0 + \beta_1 X_i + \beta_2 W_j + \beta_3 Z_{cst} + \delta Match_{ij} + \epsilon_{ijcst}, \tag{1}$$

where X, W, and Z are vectors of observed student (i), instructor (j), and course-section (cs) characteristics, respectively; t indexes semesters; $Match$ is a vector of variables that measure

[8] The Legum Baccalaureus (LLB) is an undergraduate degree in law. The Juris Doctor (JD) is the professional doctorate degree in law. The Legum Magister (LLM) is a master degree in law that typically focuses on a specialized legal area. A Ph.D. is a non-legal doctorate of philosophy. Those pursuing the LLM degree are required to have first obtained either an LLB or JD degree.

the degree of demographic similarity between student and instructor; and ϵ represents the unobserved determinants of y. We operationalize $Match$ in various ways, such as a set of four mutually exclusive race-by-sex indicators (i.e., same race and other sex, same sex and other race, same race and same sex, other race and other sex) and simpler definitions that include binary indicators for other sex and/or other race.

Of course, OLS estimates of equation (1) might be biased for several reasons. For example, despite the rich measures of student ability contained in the administrative data (e.g., LSAT scores), unobserved student characteristics might jointly predict outcomes and assignment to an other-race teacher. Similarly, equation (1) fails to control for unobserved instructor attributes, such as grading policies or teaching style. Accordingly, we follow Fairlie, Hoffmann and Oreopoulos (2014) and augment equation (1) to condition on both student and classroom fixed effects (FE), which yields our preferred specification:

$$y_{ik} = \theta_i + \omega_k + \delta Match_{ik} + \epsilon_{ik}. \tag{2}$$

Several aspects of equation (2) merit attention. First, the vectors X, W, and Z fall out of the model because they are colinear with the FE. Second, we collapse the subscripts $jcst$ into a single k subscript because identification now comes from within-classroom variation in $Match$ and classrooms are instructor, course, section, and semester specific: the classroom FE (ω) subsumes instructor, course, semester, and year FE. Third, equation (2) is only identified for outcomes that vary within-students across courses, such as course grades, due to the student FE (θ). We discuss an instrumental-variable strategy for identifying the impact of student-instructor demographic match on long-run, student-specific outcomes such as graduation in Section 4.4. Finally, there is a possible sample selection issue for the analyses of course grades, since grades are only observed for students who complete the course, and it is possible that student-instructor demographic mismatch affects the likelihood that students complete the course. This turns out to be a practically unimportant concern, as dropping courses is quite rare (occurs in only 0.6% of cases) and we find no evidence that demographic mismatch affects course dropouts.[9] We estimate equation (2) using the estimation routine proposed by Correia (2015) and compute two-way cluster-robust standard errors, which allow for correlated ϵ both within instructors across semesters and within students across courses (Cameron, Gelbach and Miller, 2012).

[9]This is perhaps unsurprising, as we are investigating required first-year courses.

3.2 Sorting Test

While the two-way FE in equation (2) address many threats to validity, one potential threat remains: differential sorting by student race or gender (Fairlie, Hoffmann and Oreopoulos, 2014). For example, the student FE controls for scenarios in which high-ability students sort into female-taught courses, but does not adequately control for gender-specific sorting processes in which high-ability female students sort into female-taught courses and high-ability male students sort into male-taught courses. To discern the extent to which differential sorting on unobservables occurs, we follow Fairlie, Hoffmann and Oreopoulos (2014) in implementing a formal test for differential sorting on observables. The test relies on the intuition of difference-in-differences estimators and the bounding procedure of Altonji, Elder and Taber (2005). It is best illustrated via an example. Suppose we want to test for differential sorting by gender. We would first compute the mean of observed student characteristic L (e.g., LSAT score) in classroom k for each gender g: $\overline{L}_k^g$. Then estimate the linear regression

$$\overline{L}_k^g = \gamma_0 + \gamma_1 Female_k + \gamma_2 1\{Female = g\} + \gamma_3 Female_k \times 1\{Female = g\}, \qquad (3)$$

where *Female* is a binary indicator equal to one if the section-k teacher is female, and zero otherwise; $1\{\cdot\}$ is the indicator function; and γ_3 is the parameter of interest. Specifically, γ_3 represents "the difference-in-differences estimate" of the average difference in observed characteristics between female and male students in female- and male-taught courses. If γ_3 is significantly different from zero, there are differences by student sex in sorting into courses on observables that systematically vary with the sex of the instructor. Alternatively, if the OLS estimate of γ_3 in equation (3) is statistically indistinguishable from zero, there is no evidence of differential sorting on observables, and thus differential sorting on unobservables in a way that would bias the two-way FE estimates of equation (2) is unlikely.

3.3 Cross-Semester Effects in Two-Course Sequences

Finally, we consider whether exposure to an other-race or other-sex instructor in the first course of a two-course sequence affects performance in the second course. Naturally, this analysis can only be conducted for the subset of first-year courses that are part of a required two-course sequence.[10] While this question can be addressed using the baseline two-way FE model given in equation (2), it is also possible to further increase the estimates' validity by

[10]There are four such sequences: Civil Procedure I & II, Legal Rhetoric I & II, Criminal Law & Criminal Procedure, and Property Law I & II.

augmenting equation (2) to condition on a second-semester course FE (φ).[11] Specifically, we estimate three-way FE models of the form

$$y_{is2} = \theta_i + \omega_{s1}^{(i)} + \varphi_{s2}^{(i)} + \delta Match_{is1} + \epsilon_{is}, \tag{4}$$

where 1 and 2 index semesters and s indexes subjects. Estimates of δ in equation (4) are robust to excluding the second-semester course FE, which is reassuring because it suggests that the demographic background of the first-semester instructor does not affect second-semester classroom assignments. Estimates of equation (4) report standard errors clustered along three dimensions: student, semester 1 instructor, and semester 2 instructor.

4 Results

This section presents the empirical results. Section 4.1 presents estimates of the sorting test characterized by equation (3). Section 4.2 presents the baseline two-way FE estimates. Section 4.3 tests for heterogeneous impacts of student-instructor demographic mismatch. Finally, Section 4.4 introduces the IV strategy for estimating the impact of faculty representation on student persistence and graduation rates.

4.1 Sorting Test Estimates

Table 3 presents estimates of the sorting test characterized by equation (3).[12] Panels A and B report estimates for differential sorting by race, comparing the average characteristics of whites and nonwhites. The former uses the entire analytic sample of first-year courses and the latter restricts the sample to second-semester courses, as students may exhibit greater selection in the second semester of their first year. Panels C and D do the same for differential sorting by sex, comparing the average characteristics of males and females.

We perform the sorting test for six outcomes: LSAT score, undergraduate GPA, median income in student's home zip code, percent of population with college degree in student's home zip code, a binary indicator equal to one if the student came from the surrounding tri-state area, and student age.[13] The LSAT and undergraduate GPA variables likely measure a combination of students' cognitive and noncognitive skills (Heckman and Kautz, 2012).

[11] This is similar to the identification strategy used by Figlio, Schapiro and Soter (2015) to identify the impact of adjunct instructors, though in that case the first-semester course FE were not included because adjunct status varies only at the classroom level.

[12] The sorting test estimates remain essentially unchanged when course-name and year FE are added to the regression.

[13] Data on undergraduate GPA is missing for the majority of students.

The zip-code information proxies for the student's socioeconomic background, which is an important predictor of undergraduate college success (Bailey and Dynarski, 2011). The "tri-state" indicator provides a crude measure of students' distances from home, which is known to predict undergraduate enrollments (Alm and Winters, 2009; Cooke and Boyle, 2011).

In Panels A and B of Table 3, 10 of 12 estimates of γ_3 are statistically insignificant, which suggests that there is little differential sorting on observables by race. Moreover, given the multiple hypotheses tested in Table 3, it is possible that the two significant results Panel A are spurious: indeed, they lose their statistical significance after adjusting for multiple comparisons (Schochet, 2009). Moreover, this result suggests sorting in the "wrong" direction in the sense that nonwhite students assigned to nonwhite faculty are from *lower* socioeconomic backgrounds, which would bias *against* finding a positive impact of demographic match on student outcomes. In Panels C and D of Table 3, all 12 estimates of γ_3 are statistically indistinguishable from zero, suggesting that there is no systematic sorting on observables by sex. In sum, the general lack of sorting on observables observed in Table 3 suggests that differential sorting on unobservables is unlikely to bias the preferred two-way FE estimates of equation (2). That the tests find no evidence of sorting in the second semester (Panels B and D), when students are arguably more able to strategically select courses, lends further credibility to a causal interpretation of the baseline estimates.

4.2 Main Results

Table 4 reports baseline estimates of equation (2) using a simple definition of *Other*: binary indicators for whether or not the student had an other-sex and other-race instructor. The first three columns of Table 4 use different definitions of the course grade as the outome. Column 1 uses a continuous measure of the course grade, which is measured on a 0-4 scale. Having an other-sex and other-race teacher significantly reduced the student's course grade by 0.02 and 0.04, respectively, though these estimates are not significantly different from one another. These effects represent small ($\approx 1\%$) changes from the average course grade of 3.3. While small in magnitude, recall that these are course-specific effects that might add up to nontrivial differences in cumulative GPA that preclude under-represented students from prestigious internships or alter the class rankings in ways that affect initial job placements starting salaries. Additionally, these small effects could be due to the effect of student-instructor demographic mismatch operating on particular margins of the course-grade distribution. Accordingly, in columns 2 and 3 we estimate linear probability models in which the outcomes are binary indicators for "good" and "bad" grades, respectively. Consistent with the results in column 1, column 2 shows a significant, negative effect of demographic mismatch on the probability that students earn an $A-$ or A. The racial mismatch effect is larger than the sex mismatch effect, but the difference is not statistically significant. These effects are arguably economically significant, as the other-race effect of 0.03 constitutes 9% of the sample average "good-grade" rate. Column 3 shows that there is no effect of student-instructor demographic mismatch on the likelihood of receiving a "bad grade" ($< B-$). The remaining columns of Table 4 show that there are neither effects of mismatch on the likelihood that the student takes another course in the subject nor on the likelihood that the student drops the course.[14] The latter null result is important, however, as it suggests that the sample selection inherent in the course-grade analyses is negligible.

It is also possible that there are multiplicative effects of having both an other-race *and* other-sex instructor. Table 5 investigates this possibility by specifying *Other* as a set of four mutually-exclusive categorical indicators, with same-sex *and* same-race as the omitted reference category. The course-grade results reported in columns 1-3 are broadly consistent with those in Table 4: there are negative effects of student-instructor demographic mismatch on the probability of receiving a "good grade," which are larger for racial-mismatch than for sex-mismatch, but not on the probability of receiving a "bad grade." Interestingly, however, while the impact of having both an other-race and other-sex instructor is larger than that of having an other-race but same-sex instructor, the difference is not statistically significant.

Finally, Table 6 reports estimates of equation (4), which show the impact of having an other-sex or other-race instructor in the first course of a required two-course sequence on performance in the second course. Two versions of equation (4), with and without second-semester classroom FE, are estimated for each of three outcomes: course grade, a binary indicator for "good grade" (i.e., A or A-), and a binary indicator for "bad grade" (i.e., <B-). For each outcome, the point estimates are robust to excluding the second-semester course FE, which suggests that the demographic background of the first-semester instructor does not affect second-semester classroom assignments. This is to be expected, as students generally do not select specific sections in their first year, but is reassuring nonetheless.

The results are broadly similar to the baseline two-way FE estimates reported in Table 4: there are negative effects of student-instructor mismatch in the first course on grades and on the probability of receiving a good grade in the second course. Once again, the other-race effect is about twice as large as the other-sex effect, though here only the other-sex effect is statistically significant at traditional confidence levels. That these estimates are qualitatively similar to the baseline estimates reported in Table 4 lends further credence to a causal interpretation of the relationship between student-instructor demographic mismatch

[14]The sample size for subsequent course taking is smaller because there are not subsequent courses in all required first-year courses.

and course grades. Moreover, this similarity sheds some light on the mechanisms at work, as the cross-semester effects documented in Table 6 must be due to students' responses to the instructor, rather than a purely instructor-based channel, such as biased grading.

4.3 Heterogeneity

In this section we test for heterogeneity in the impacts of student-instructor demographic mismatch on the probability of receiving a "good grade" documented in section 4.2. Column 1 of Table 7 estimates an augmented version of the baseline model that allows the other-race and other-sex effects to vary by race, since past research on the impact of student-instructor mismatch finds larger effects among racial-minority students (Dee, 2004). While none of the individual coefficients are statistically significant, the net effects of having an other-sex and other-race instructor are statistically significant for nonwhite students, which suggests that the main results were driven by the responses of nonwhite students to demographic mismatch in the classroom. Moreover, the net effects of -0.02 and -0.04 are arguably economically significant, as they represent effects of 9% and 17%, respectively, from the nonwhite mean of 0.23.

The impacts of student-teacher demographic mismatch might also vary by student sex. Unfortunately, the classroom-FE identification strategy does not allow for a female-other sex interaction term analogous to the nonwhite-other race interaction term included in column 1. However, we can interact the other-race indicator and the nonwhite interactions with a female indicator, and we report these estimates in column 2 of Table 7. Like in column 1, the individual coefficient estimates are imprecise. However, the net effects of having an other-sex and other-race instructor for nonwhite females are both marginally statistically significant, and relatively large in magnitude: they constitute approximately 12% and 22% effects, respectively. Together, these results suggest that the impact of representation among law school instructors is greatest for nonwhite students, particularly for nonwhite females.

We find no evidence of heterogeneity along other observable student dimensions, such as students' ability (LSAT score), age, home region, and zip-code SES. Nor do we find evidence of heterogeneity by observable instructor characteristics, such as experience, rank of JD program, or faculty rank (i.e., adjunct, teaching-track, tenure-line, tenured). These null results are not reported in the interest of brevity.

Finally, we test for heterogeneity in the impact of student-instructor demographic mismatch by classroom characteristics, as classroom environments might moderate the impact of mismatch (Ho and Kelman, 2014; Inzlicht and Ben-Zeev, 2000). First, we allow the effect of mismatch to vary by class size. Whether larger classrooms magnify or dampen the mismatch effects documented previously is theoretically ambiguous, as smaller classrooms might shine a spotlight on implicit biases, or facilitate relationships that supercede stereotypes. Second, we allow the effect of mismatch to vary with the demographic composition of classrooms, as the impact of an other-race or other-sex instructor might be more pronounced in less diverse settings in which female or nonwhite students feel isolated. Given the exploratory nature of this analysis, we model the heterogeneity using quadratics in class size and percent female (nonwhite).

Appendix Table A1 shows that the quadratic terms are at least marginally jointly significant in each specification, though for ease of interpretation we plot the marginal effects as functions of class size and percent female (nonwhite). Specifically, Figure 1 plots the marginal effects on the probability of receiving an A/A- of having an other-race or other-sex instructor as a function of class size. Interestingly, there is essentially no effect of mismatch in the smallest classes. The other-sex effect monotonically increases in magnitude with class size, though at a relatively slow pace. The other-race effect, meanwhile, exhibits a U-shaped pattern. The deleterious effect of having an other-race instructor is largest in classrooms of about 60 students. One possible interpretation of this pattern is that the personal connections and relative anonymity in very small and very large classes, respectively, mitigate the harm associated with having an other-race instructor.

Similarly, Figure 2 plots the marginal effects on the probability of receiving an A/A- of having an other-race (other-sex) instructor as a function of the fraction of the classroom that is female (nonwhite). The other-race effect is fairly constant at about -0.03 or -0.04, regardless of the proportion of nonwhite students in the class. The effect only approaches zero when the fraction nonwhite approaches one, which is an out-of-sample prediction. However, the other-sex effect is highly nonlinear. Intuitively, it is most pronounced when females comprise less than 40% of the class. The other-sex effect approaches zero when 60 to 70% of the class is female. This is suggestive of stereotype threat (Steele, 1997), whereby females disengage with law school when they perceive themselves as outsiders, and consistent with experimental evidence that shows that the gender ratio of a classroom affects female's test performance, but not male's (Inzlicht and Ben-Zeev, 2000).

4.4 Long-Run Outcomes

The results presented thus far provide robust, arguably causal evidence that having an other-race or other-gender law school instructor in a first-year required course significantly reduces the likelihood of receiving a "good grade" in that course, particularly for non-white and non-white female students. This suggests that law school faculty representation affects the intensive margin of student success, as course grades are likely valued in the lawyer labor market (Oyer and Schaefer, 2016). However, the two-way fixed effects strategy used to identify these effects cannot be used to examine the effects of representation on persistence and graduation rates, because there is no within-student variation in these extensive-margin outcomes. To examine whether demographic representation in first-year law courses affects students' persistence and eventual graduation, we use an instrumental variables (IV) strategy similar to that used in Bettinger and Long (2005, 2010) and Fairlie, Hoffmann and Oreopoulos (2014).

The goal is to estimate student-level linear probability models (LPM) of the form

$$Pr(y_i = 1) = \tau FractionMatched_{it} + \beta X_i + \nu_{it}, \tag{5}$$

where $FractionMatched$ is simply the fraction of student-i's first-year instructors of the same race or sex as the student and X is a vector of observed student socio-demographic and pre-admission ability measures (i.e., LSAT scores). Even with rich controls in X, OLS estimates of equation (5) are potentially biased by unobserved student factors that jointly predict $FractionMatched$ and y.

Accordingly, we estimate equation (5) separately for female (nonwhite) students by 2SLS, using cohort-specific deviations from the steady-state level of first-year female (nonwhite) instructors to instrument for $FractionMatched$. This IV strategy exploits arguably random variation over time in the demographic composition of the first-year core faculty attributable to retirements, sabbaticals, adjunct hires, and so on (Bettinger and Long, 2005, 2010). Specifically, we create cohort-specific instruments that are the difference between the share of female (nonwhite) first-year core instructors in the student's first year and the share of female (nonwhite) first-year core instructors over what would be the student's first four years (133% of expected time to degree) in the program.[15] Figure 3 documents significant variation in the instruments across cohorts.

Table 8 presents 2SLS estimates of equation (5) alongside the corresponding first-stage estimates. Columns (1)-(3) do so for the sample of female students, where $FractionMatched$ and its corresponding instrument measure the fraction of first-year courses taught by female instructors. Moving from left to right, columns (1)-(3) report estimates of models that take richer specifications of X, which both increase precision of the 2SLS estimates and account for the composition of incoming classes. Panel A reports the first-stage coefficients, which are uniformly strong. Panel B reports 2SLS estimates for females' persistence into the second year. There is no evidence that exposure to female faculty in the first year increases

[15] The IV results are robust to slight changes in the construction of the instrument.

the likelihood that female students persist into their second year of law school. Panel C of Table 8 presents 2SLS estimates for the probability that females graduate within three years. Here, there is modest evidence of an impact of female faculty representation on the timely graduation of female students.

Columns (4)-(6) do the same for the sample of nonwhite students. The first-stages are again strong and the 2SLS estimates in Panels B and C suggest a statistically significant effect of cumulative exposure to nonwhite faculty in first-year courses on the probability that nonwhite students persist into the fall of their second year of law school and graduate with a JD within three years, respectively. Specifically, these estimates suggest that a ten percentage point increase in the fraction of first-year courses (roughly one course) taught by nonwhite instructors increases the probability of persisting into year two by about 5 to 10 percentage points.

5 Conclusion

We use rich student-instructor matched administrative data from a large, private, top-100 law school to provide novel evidence on the causal relationship between student-instructor demographic match and student outcomes in the law school context. Two-way student and course fixed-effects models provide arguably causal estimates of the impact of such mismatch on short-run (course-specific) outcomes such as course grades. Sorting tests provide no evidence of endogenous sorting on observables into classrooms, which suggests that endogenous sorting on unobservables is unlikely to bias the baseline two-way FE estimates.

These preferred estimates suggest that having an other-race or other-sex instructor in a first-year required course significantly reduces the likelihood of earning an A/A- in the course. Specifically, having an other-sex instructor reduces the likelihood of receiving a good grade (A or A-) by one percentage point (3%) and having an other-race instructor reduces the likelihood of receiving a good grade by three percentage points (10%). This result has the potential to contribute to pay gaps, as Oyer and Schaefer (2016) provide descriptive evidence of a wage-class rank gradient in law schools outside the top 10.[16] However, we find no effects of student-instructor demographic mismatch on dropping courses or taking subsequent courses in the same field, nor do we find effects at other points of the grade distribution. Consistent with previous research in the K-12 context, these effects are stronger for nonwhite students, especially nonwhite females. It is actually somewhat remarkable that

[16]Our own analyses of the publicly available *After the JD* survey data confirm the positive association between law school GPA and earnings both overall, and for specific demographic groups, for lawyers who attended non-top 10 law schools. See Appendix B for details.

the same harmful effects of mismatch observed among relatively vulnerable populations of primary school, community college, and first-year college students are observed in an elite law school setting, given the successes and experiences necessary for admission to a top-100 law school. Indeed, a broad takeaway of the current study is that the race- and sex-interactions associated with stereotype threat likely pervade society in substantive, if unexpected ways.

These results suggest that diversity in the legal profession, and the status of women and people of color in the legal profession, would be improved by increasing the diversity of law school faculty. However, whether and how these results would generalize to other law schools, particularly those with less diverse student and faculty populations, remains an open question worthy of future exploration. There are also questions regarding the general equilibrium responses to the hiring of a more diverse faculty, and potential supply-side limitations of such faculty in the short run. For these reasons, another potential policy response is to provide law school (and university) faculty with theoretically-informed implicit bias training, which has proven to be effective in some early pilots (Carnes et al., 2015). Similarly, Darling-Hammong and Holmquist (2015) provide suggestions to law school faculty on how to better serve historically underrepresented students, many of which echo the theoretically-informed, "WISE" interventions and strategies advocated by social psychologists (Okonofua, Paunesku and Walton, 2016; Walton, 2014).

References

Alm, James, and John V. Winters. 2009. "Distance and Intrastate College Student Migration." *Economics of Education Review*, 28(6): 728–738.

Altonji, Joseph G., and Rebecca M. Blank. 1999. "Race and Gender in the Labor Market." *Handbook of Labor Economics*, 3: 3143–3259.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151–184.

Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik. 2015. "The Effect of Teacher Gender on Student Achievement in Primary School." *Journal of Labor Economics*, 33(1): 63–89.

Azmat, Ghazala, and Rosa Ferrer. Forthcoming. "Gender Gaps in Performance: Evidence From Young Lawyers." *Journal of Political Economy.*

Bailey, Martha J., and Susan M. Dynarski. 2011. "Gains and Gaps: Changing Inequality in US College Entry and Completion." National Bureau of Economic Research.

Banks, Taunya Lovell. 1988. "Gender Bias in the Classroom." *Journal of Legal Education*, 38(1/2): 137–146.

Bertrand, Marianne. 2011. "New Perspectives on Gender." *Handbook of Labor Economics*, 4: 1543–1590.

Bertrand, Marianne, and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *Industrial & Labor Relations Review*, 55(1): 3–21.

Bettinger, Eric P., and Bridget Terry Long. 2005. "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students." *The American Economic Review*, 95(2): 152–157.

Bettinger, Eric P., and Bridget Terry Long. 2010. "Does cheaper mean better? The impact of using adjunct instructors on student outcomes." *The Review of Economics and Statistics*, 92(3): 598–613.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2012. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Carnes, Molly, Patricia G. Devine, Linda Baier Manwell, Angela Byars-Winston, Eve Fine, Cecilia E. Ford, Patrick Forscher, Carol Isaac, Anna Kaatz, Wairimu Magua, et al. 2015. "Effect of an intervention to break the gender bias habit for faculty at one institution: a cluster randomized, controlled trial." *Academic Medicine: Journal of the Association of American Medical Colleges*, 90(2): 221.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *The Quarterly Journal of Economics*, 125(3): 1101–1144.

Cooke, Thomas J., and Paul Boyle. 2011. "The Migration of High School Graduates to College." *Educational Evaluation and Policy Analysis*, 33(2): 202–213.

Correia, Sergio. 2015. "REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-Dimensional Fixed Effects." *Statistical Software Components*.

Darling-Hammong, Sean, and Kristen Holmquist. 2015. "Creating Wise Classrooms to Empower Diverse Law Students: Lessons in Pedagogy from Transformative Law Professors." *Chicana/o-Latina/o L. Rev.*, 33: 1.

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

Dee, Thomas S. 2005. "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *The American Economic Review*, 95(2): 158–165.

Dee, Thomas S. 2007. "Teachers and the Gender Gaps in Student Achievement." *Journal of Human Resources*, 42(3): 528–554.

Delisle, Jason. 2014. "The Graduate Student Debt Review: The State of Graduate Student Borrowing." *New America Education Policy Program*.

Dinovitzer, Ronit, Nancy Reichman, and Joyce Sterling. 2009. "The Differential Valuation of Women's Work: A New Look at the Gender Gap in Lawyers' Incomes." *Social Forces*, 88(2): 819–864.

Eren, Ozkan, and Naci Mocan. 2016. "Emotional Judges and Unlucky Juveniles." National Bureau of Economic Research.

Fairlie, Robert W., Florian Hoffmann, and Philip Oreopoulos. 2014. "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." *The American Economic Review*, 104(8): 2567–2591.

Figlio, David N., Morton O. Schapiro, and Kevin B. Soter. 2015. "Are tenure track professors better teachers?" *Review of Economics and Statistics*, 97(4): 715–724.

Gershenson, Seth, Stephen B. Holt, and Nicholas W. Papageorge. 2016. "Who Believes in Me? The Effect of Student–Teacher Demographic Match on Teacher Expectations." *Economics of Education Review*, 52: 209–224.

Guinier, Lani, Michelle Fine, Jane Balin, Ann Bartow, and Deborah Lee Stachel. 1994. "Becoming Gentlemen: Women's Experiences at One Ivy League Law School." *University of Pennsylvania Law Review*, 143(1): 1–110.

Heckman, James J., and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics*, 19(4): 451–464.

Ho, Daniel E., and Mark G. Kelman. 2014. "Does class size affect the gender gap? A natural experiment in law." *The Journal of Legal Studies*, 43(2): 291–321.

Hoffmann, Florian, and Philip Oreopoulos. 2009. "A Professor Like Me: The Influence of Instructor Gender on College Achievement." *Journal of Human Resources*, 44(2): 479–494.

Holder Jr, Eric H. 2001. "The Importance of Diversity in the Legal Profession." *Cardozo L. Rev.*, 23: 2241.

Inzlicht, Michael, and Talia Ben-Zeev. 2000. "A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males." *Psychological Science*, 11(5): 365–371.

Loury, Glenn C. 2009. *The Anatomy of Racial Inequality.* Harvard University Press.

Lusher, Lester, Doug Campbell, and Scott Carrell. 2015. "TAs Like Me: Racial Interactions between Graduate Teaching Assistants and Undergraduates." National Bureau of Economic Research.

Mustard, David B. 2001. "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence From the US Federal Courts." *Journal of Law and Economics*, 44(1): 285–314.

Okonofua, Jason A., David Paunesku, and Gregory M. Walton. 2016. "Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents." *Proceedings of the National Academy of Sciences*, 113(19): 5221–5226.

Oyer, Paul, and Scott Schaefer. 2015. "Firm/Employee Matching: An Industry Study of US Lawyers." *Industrial & Labor Relations Review*, 69(2): 378–404.

Oyer, Paul, and Scott Schaefer. 2016. "The Returns to Elite Degrees: The Case of American Lawyers." *Unpublished Manuscript*.

Papageorge, Nicholas W., Seth Gershenson, and Kyungmin Kang. 2016. "Teacher Expectations Matter." IZA Discussion Paper No. 10165.

Rachlinski, Jeffrey J., Sheri Lynn Johnson, Andrew J. Wistrich, and Chris Guthrie. 2008. "Does Unconscious Racial Bias Affect Trial Judges." *Notre Dame L. Rev.*, 84: 1195.

Schochet, Peter Z. 2009. "An Approach for Addressing the Multiple Testing Problem in Social Policy Impact Evaluations." *Evaluation Review*, 33(6): 539–567.

Schwarcz, Daniel, and Dion Farganis. Forthcoming. "The Impact of Individualized Feedback on Law Student Performance." *Journal of Legal Education*.

Steele, Claude M. 1997. "A threat in the air: How stereotypes shape intellectual identity and performance." *American Psychologist*, 52(6): 613.

Walton, Gregory M. 2014. "The new science of wise psychological interventions." *Current Directions in Psychological Science*, 23(1): 73–82.

Wilkins, David B., and G. Mitu Gulati. 1996. "Why Are There So Few Black Lawyers in Corporate Law Firms? An Institutional Analysis." *California Law Review*, 493–625.

Wood, Robert G., Mary E. Corcoran, and Paul N. Courant. 1993. "Pay Differences Among the Highly Paid: The Male-Female Earnings Gap in Lawyers' Salaries." *Journal of Labor Economics*, 11(3): 417–441.

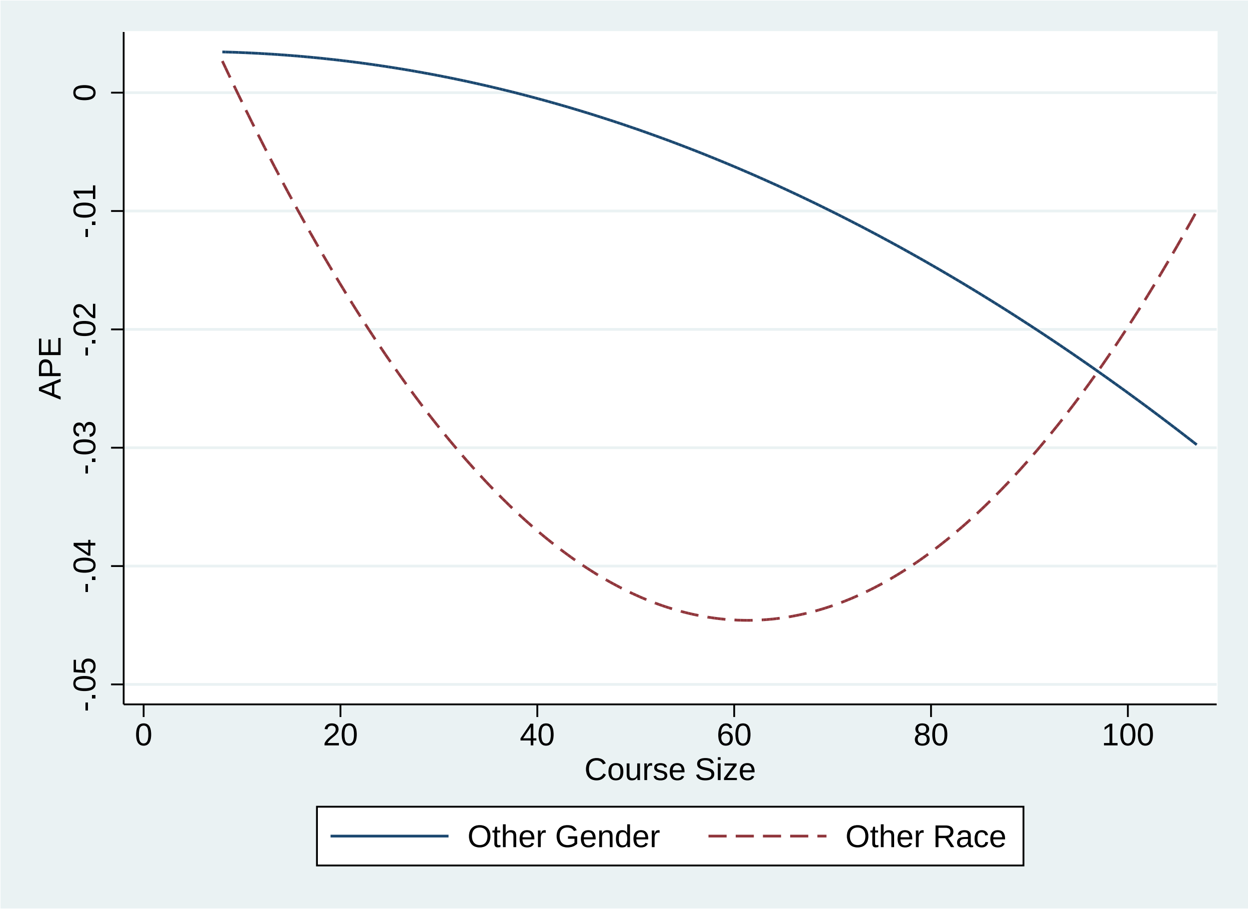

Figure 1: Average Partial Effects (APE) of Student-Instructor Mismatch on the Probability of Receiving a “Good Grade” as a Function of Class Size

“Good Grade” is defined as an A or A-.

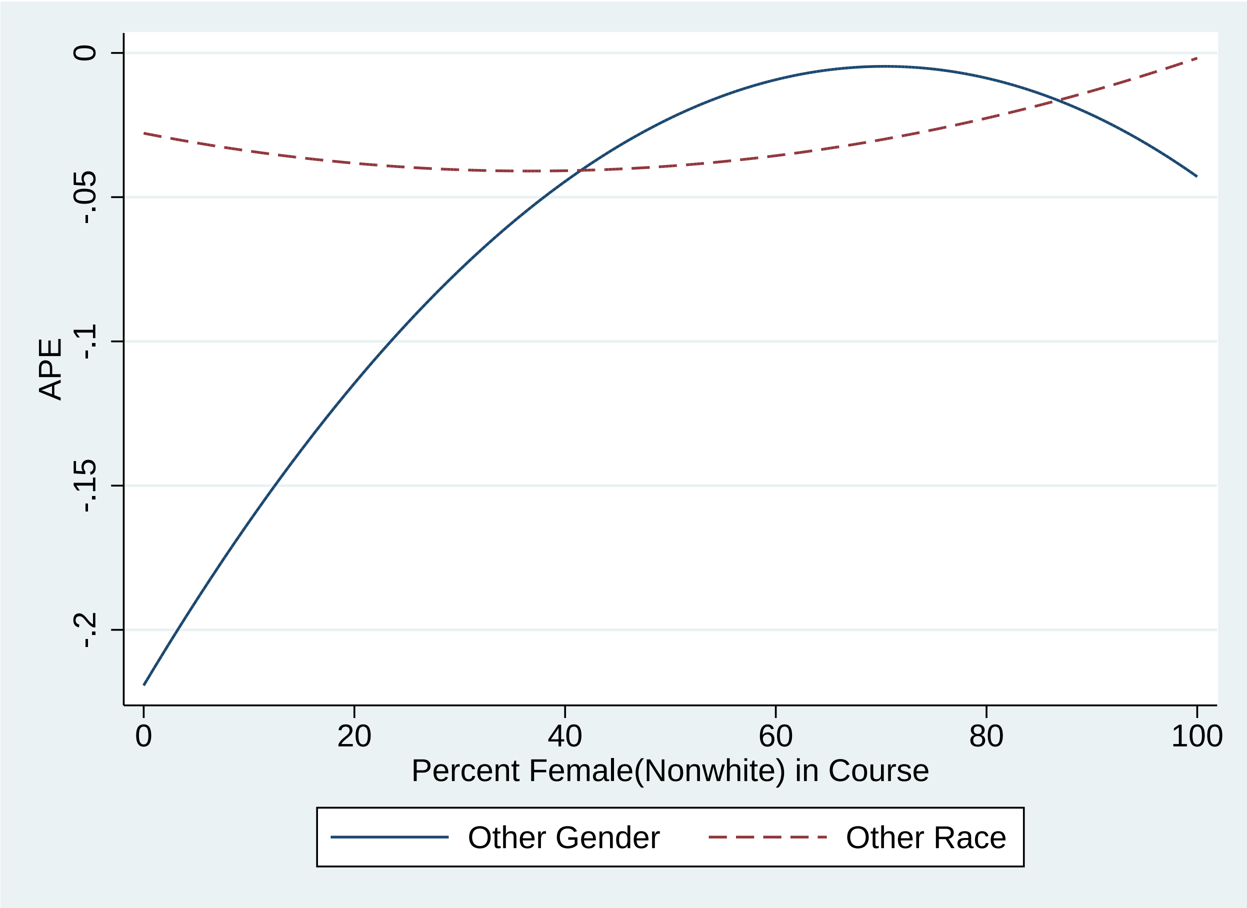

Figure 2: Average Partial Effects (APE) of Student-Instructor Mismatch on the Probability of Receiving a "Good Grade" as a Function of Class Composition

"Good Grade" is defined as an A or A-.

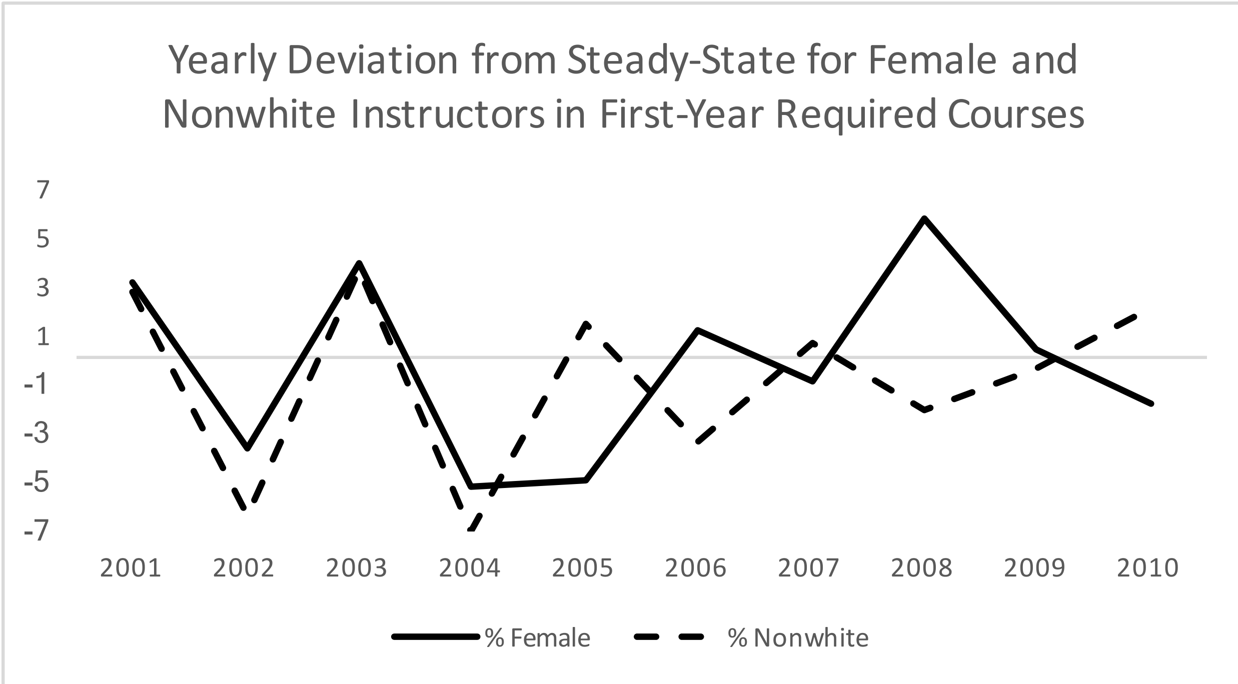

Figure 3: Deviations from Steady State in First-Year Core Faculty Demographics

Note: Zero deviation represents the steady state.

Table 1: Sample Statistics for Students and Instructors in First-Year Required Courses

	White		Non-white		Male		Female	
Panel A: Student-Course Level	mean	sd	mean	sd	mean	sd	mean	sd
Course Grade (0-4)	3.36	0.46	3.15	0.51	3.27	0.50	3.29	0.49
Take Another Course	0.80		0.81		0.80		0.81	
Take Another Property Course	0.79		0.81		0.82		0.78	
Take Another Litigation Course	0.87		0.88		0.86		0.88	
Take Another Constitutional Course	0.70		0.69		0.68		0.71	
Take Another Criminal Course	0.85		0.87		0.84		0.87	
Dropped Course	0.004		0.004		0.004		0.004	
Grade: A	0.40		0.23		0.33		0.35	
Grade: B	0.56		0.67		0.60		0.60	
Grade: C, D, F	0.04		0.10		0.06		0.06	
Other Gender	0.51		0.52		0.42		0.58	
Other Race	0.18		0.95		0.41		0.50	
Observations	23,532		13,510		15,461		21,581	
Panel B: Student Level								
Age (First Semester)	25.60	2.63	25.40	2.45	25.80	2.67	25.40	2.48
Female Student	0.54		0.65		0.00		1.00	
Black Student	0.00		0.21		0.05		0.10	
Latino Student	0.00		0.37		0.12		0.14	
Asian Student	0.00		0.35		0.10		0.14	
White Student	1.00		0.00		0.70		0.59	
Other Student	0.00		0.08		0.03		0.03	
LSAT	161.10	3.38	156.20	4.85	159.90	4.38	158.80	4.80
Persist to Second Year	0.90		0.92		0.89		0.91	
Graduated	0.71		0.71		0.70		0.72	
Observations	3,018		1,736		1,993		2,761	
Panel C: Instructor Level								
Non-white Instructor	0.00		1.00		0.13		0.21	
Black Instructor	0.00		0.49		0.07		0.10	
Latino Instructor	0.00		0.23		0.03		0.05	
Asian Instructor	0.00		0.29		0.04		0.06	
White Instructor	1.00		0.00		0.87		0.79	
Female Instructor	0.46		0.60		0.00		1.00	
Years of AU Experience	5.87	9.62	2.86	5.62	7.71	11.00		5.95
JD	0.94		0.97		0.93		0.97	
JD Rank	37.90	36.00	37.10	42.60	36.40	39.40	39.20	34.60
PhD	0.10		0.06		0.09		0.09	
LLM	0.09		0.18		0.13		0.07	
LLB	0.04		0.00		0.06		0.01	
Observations	171		35		106		100	

Table 2: Sample Statistics for First-Year Required Courses

Course Level Characteristics	mean	sd	Course Names	Pct
Class Size	42.90	33.30	501 - Civil Procedure	5.39
Female Students	0.59	0.15	502 - Civil Procedure II	2.12
Age (First Semester)	25.90	2.30	503 - Constitutional Law	6.10
Black Students	0.07	0.07	504 - Contracts	5.65
Latino Students	0.13	0.12	507 - Criminal Law	6.10
Asian Students	0.13	0.11	508 - Criminal Procedure I	2.74
White Students	0.64	0.15	516 - Legal Rhetoric I	26.33
Other Students	0.03	0.05	517 - Legal Rhetoric II	28.80
Female Instructor	0.45	0.50	518 - Property	5.57
Black Instructor	0.08	0.27	519 - Property II	1.77
Asian Instructor	0.02	0.15	522 - Torts	5.04
Latino Instructor	0.04	0.18	550 - Legal Ethics	4.42
White Instructor	0.86	0.35	Observations	1,132
More than one instructor race choice term	0.74	0.44		
More than one instructor race choice year	0.76	0.43		
More than one instructor gender choice term	0.94	0.24		
More than one instructor gender choice year	0.95	0.21		
Observations	1,132			

Notes: Classroom level demographics are presented as proportions.

Table 3: Sorting Test Estimates

			Outcome			
	LSAT	UGPA	Median Income (Zip)	% Adult w/ BA (Zip)	In/Nearby State	Student Age
Average:	159.27	3.42	77151.54	37.19	0.52	25.5
Panel A: Sorting by Race - Semester 1 & 2						
Nonwhite Instructor	-0.07	-0.04	-2824.01**	4.08***	-0.01	-0.00
	(0.29)	(0.06)	(1344.09)	(1.32)	(0.02)	(0.07)
Nonwhite Student	-4.59***	-0.15***	-2547.76***	-2.19***	-0.02**	-0.11**
	(0.12)	(0.03)	(614.05)	(0.29)	(0.01)	(0.05)
Nonwhite Instructor * Nonwhite Student	**0.38**	**-0.04**	**-1932.42**	**-2.35****	**0.01**	**0.19***
	(0.29)	(0.10)	(1371.18)	(0.95)	(0.03)	(0.11)
Constant	160.41***	3.46***	78669.82***	38.01***	0.53***	25.60***
	(0.10)	(0.02)	(547.63)	(0.45)	(0.01)	(0.03)
Observations	1952	488	2152	2152	2152	2152
Panel B: Sorting by Race - Semester 2						
Nonwhite Instructor	0.14	-0.03	-260.29	1.27	-0.03*	0.01
	(0.39)	(0.08)	(1783.34)	(1.67)	(0.02)	(0.09)
Nonwhite Student	-4.45***	-0.14***	-2342.26**	-2.13***	-0.02	-0.10
	(0.17)	(0.05)	(917.78)	(0.42)	(0.01)	(0.07)
Nonwhite Instructor * Nonwhite Student	**0.43**	**-0.05**	**-717.22**	**-1.34**	**0.03**	**0.16**
	(0.39)	(0.13)	(1713.99)	(0.93)	(0.03)	(0.14)
Constant	160.29***	3.46***	78415.64***	38.19***	0.53***	25.57***
	(0.15)	(0.03)	(797.32)	(0.63)	(0.01)	(0.05)
Observations	1017	244	1104	1104	1104	1104
Panel C: Sorting by Gender - Semester 1 & 2						
Female Instructor	0.17	-0.06	-635.80	0.77	0.01	-0.09
	(0.19)	(0.06)	(1161.88)	(0.80)	(0.01)	(0.07)
Female Student	-1.02***	0.11**	-1828.67**	0.34	0.02**	-0.36***
	(0.12)	(0.04)	(724.39)	(0.32)	(0.01)	(0.07)
Female Instructor * Female Student	**0.02**	**-0.00**	**-659.00**	**-0.29**	**-0.00**	**-0.00**
	(0.19)	(0.07)	(1213.58)	(0.55)	(0.02)	(0.10)
Constant	159.20***	3.36***	78496.58***	36.89***	0.50***	25.84***
	(0.13)	(0.04)	(753.45)	(0.54)	(0.01)	(0.05)
Observations	1963	466	2186	2186	2192	2192
Panel D: Sorting by Gender - Semester 2						
Female Instructor	-0.02	-0.07	10.83	-0.49	0.01	-0.00
	(0.26)	(0.08)	(1641.06)	(1.11)	(0.02)	(0.10)
Female Student	-0.86***	0.11*	-1683.97*	0.48	0.03	-0.35***
	(0.17)	(0.06)	(1009.59)	(0.47)	(0.02)	(0.09)
Female Instructor * Female Student	**-0.04**	**0.01**	**-707.00**	**-0.53**	**0.00**	**0.04**
	(0.26)	(0.10)	(1718.89)	(0.78)	(0.03)	(0.14)
Constant	159.19***	3.36***	78639.15***	37.46***	0.51***	25.78***
	(0.18)	(0.05)	(1069.76)	(0.77)	(0.01)	(0.07)
Observations	1030	234	1137	1137	1142	1142

Note: Each column represents tests for sorting on a different student background characteristic. UGPA is undergraduate grade point average.In/Nearby State is a binary variable indicating the student's home address is within the same state as the institution or a bordering state.
Standard errors in parentheses are clustered by course.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Impact of demographic mismatch on first-year required course outcomes.

	(1) Continuous Grade	(2) A/A- Grade	(3) C, D, F Grade	(4) Take Another	(5) Dropped Course
Other Gender (1)	-0.02**	-0.01**	0.00	0.02	-0.00
	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Other Race (2)	-0.04**	-0.03***	0.01	-0.00	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Differences in coefficients *(P)*					
1=2	0.21	0.20	0.26	0.18	0.26
Observations	37042	37042	37042	18748	37263
Course FE	Yes	Yes	Yes	Yes	Yes
Student FE	Yes	Yes	Yes	Yes	Yes

Note: Each column represents a different model specification. Column 4 has fewer observations because not all required courses correspond to elective course subjects. Standard errors in parentheses are clustered by student and instructor.

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

Table 5: Impact of demographic mismatch on first-year required course outcomes.

	(1) Continuous Grade	(2) A/A- Grade	(3) C, D, F Grade	(4) Take Another	(5) Dropped Course
Same Race, Mismatch Gender (1)	-0.01*	-0.01	0.00	0.03**	-0.00
	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Mismatch Race, Match Gender (2)	-0.04**	-0.03**	0.01	0.02*	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Mismatch Race, Mismatch Gender (3)	-0.05***	-0.04***	0.01	0.02	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Differences in coefficients *(P)*					
1=2	0.22	0.02	0.26	0.17	0.27
1=3	0.03	0.02	0.27	0.12	0.27
2=3	0.05	0.09	0.03	0.98	0.76
Observations	37042	37042	37042	18748	37263
Course FE	Yes	Yes	Yes	Yes	Yes
Student FE	Yes	Yes	Yes	Yes	Yes

Note: Each column represents a different model specification. Column 4 has fewer observations because not all required courses correspond to elective course subjects. Standard errors in parentheses are clustered by student and instructor.

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

Table 6: Cross-Semester Effects of Demographic Mismatch in Two-Course Sequences

	(1)	(2)	(3)	(4)	(5)	(6)
	Continuous Grade		A/A-		C, D, F Grade	
Other Gender	-0.032**	-0.032**	-0.034**	-0.034**	0.004	0.004
	(0.012)	(0.012)	(0.016)	(0.016)	(0.008)	(0.008)
Other Race	-0.079	-0.077	-0.073	-0.072	0.017	0.017
	(0.050)	(0.050)	(0.048)	(0.048)	(0.021)	(0.021)
Course 1 FE	Yes	Yes	Yes	Yes	Yes	Yes
Course 2 FE	No	Yes	No	Yes	No	Yes
Student FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: N = 4342. Each column represents a different model specification. Standard errors in parentheses are clustered three ways: by student, first instructor, and second instructor.

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

Table 7: Heterogeneous Effects of Student-Instructor Demographic Mismatch

	(1) A/A- Grade	(2) A/A- Grade
Other Gender	-0.01	-0.01
	(0.01)	(0.01)
Other Race	-0.02	-0.03
	(0.02)	(0.02)
OG * Non-white	-0.01	0.00
	(0.01)	(0.02)
OR * Non-white	-0.01	0.01
	(0.03)	(0.04)
OG * Female * Nonwhite Student		-0.02
		(0.02)
OR * Female Student		0.01
		(0.02)
OR * Female * Nonwhite Student		-0.04
		(0.05)
Net Effects		
OG Non-white APE	-0.02**	-0.01
	(0.01)	(0.01)
OG Non-white Female APE		-0.03*
		(0.02)
OR Non-white APE	-0.04*	-0.01
	(0.02)	(0.03)
OR Non-white Female APE		-0.05*
		(0.03)
Observations	37042	37042
Course FE	Yes	Yes
Student FE	Yes	Yes

Note: Each column represents a different model specification. Standard errors in parentheses are clustered by student and instructor.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Instrumental Variables (IV) Estimates of Faculty Representation on Attainment

	Female Students			Nonwhite Students		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: First Stage						
	1.31***	1.31***	1.23***	0.76***	0.76***	0.58***
	(0.10)	(0.10)	(0.11)	(0.12)	(0.12)	(0.16)
Panel B: Persist to Second Fall						
% Match	0.0003	0.0003	-0.0005	0.0052**	0.0051**	0.0111**
	(0.0010)	(0.0010)	(0.0012)	(0.0025)	(0.0025)	(0.0048)
Mean DV	0.9446	0.9446	0.9450	0.9430	0.9430	0.9457
Observations	2256	2256	1746	1421	1421	1161
Panel C: Timely Graduation (3 Years)						
% Match	0.0030**	0.0030**	0.0025	0.0060	0.0058	0.0171**
	(0.0014)	(0.0014)	(0.0017)	(0.0038)	(0.0038)	(0.0075)
Mean DV	0.8865	0.8865	0.8877	0.8712	0.8712	0.8768
Observations	2256	2256	1746	1421	1421	1161
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Student Demographics	No	Yes	Yes	No	Yes	Yes
LSAT Score	No	No	Yes	No	No	Yes

Note: Robust standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix A

Table A1: Heterogeneity by Course Size and Percent Female(Nonwhite) in Course

	Course Size		Percent Female (Nonwhite) in Course	
	(1)	(2)	(3)	(4)
Other Gender	0.003412		-0.219260	
	(0.024841)		(0.167925)	
Other Gender * Course Size	0.000029			
	(0.001040)			
Other Gender * Course Size (Sq)	-0.000003			
	(0.000009)			
Other Race		0.017907		-0.027838***
		(0.024187)		(0.010603)
Other Race * Course Size		-0.002037**		
		(0.000913)		
Other Race * Course Size (Sq)		0.000017**		
		(0.000008)		
Other Gender * Percent Female			0.006104	
			(0.005567)	
Other Gender * Percent Female (Sq)			-0.000043	
			(0.000046)	
Other Race * Percent Nonwhite				-0.000715
				(0.000798)
Other Race * Percent Nonwhite (Sq)				0.000010
				(0.000019)
P-Value for Joint Significance Tests	0.076*	0.007***	0.080*	0.029**
Observations	37042	37042	37042	37042
Course FE	Yes	Yes	Yes	Yes
Student FE	Yes	Yes	Yes	Yes

Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B

This appendix uses publicly available data from *After the JD* (AJD) to document the descriptive relationship between law school grades and early-career salaries for individuals who earned JDs from non-top 10 law schools.[17] The motivation for this appendix is to show that the impacts of student-instructor mismatch on course grades documented in the current study likely translate into demographic pay gaps among early-career law professionals.

The public-use AJD data report annual earnings in 8 bins: <\$40,000, \$40,000-\$49,999, \$50,000-\$59,999, \$60,000-\$74,999, \$75,000-\$99,999, \$100,000-\$124,999, \$125,000-\$149,999, and >\$150,000. Accordingly, we estimate descriptive ordered-logit models in which this categorical annual-earnings variable is the dependent variable. Appendix Table B1 reports the ordered-logit coefficients for the full sample. The parsimonious specifications in columns 1 and 2 document the unconditional female pay gap and wage-GPA gradient, respectively. The omitted reference category for the GPA variable is <3.0. Column 4 shows that these patterns are robust to controlling for law school quality.

Because the ordered-logit coefficients are not directly interpretable, Appendix Table B2 reports the average partial effects (APE) of these covariates on the probability of being in each earnings band for the fully-specified, full-sample estimates reported in column 4 of Appendix Table B1. Here we see that females are two to four percentage points more likely than males to be in the lowest-earning categories and two to four percentage points less likely than men to be in the highest-earning categories. The APE for the categorical GPA indicators show that each 0.25 increase in GPA is associated with a two to four percentage point increase in the probability of being in one of the high-earnings brackets, and a symmetric decrease in the probability of being in a low-earning bracket. Importantly, this suggests that even a relatively small change in GPA attributable to student-instructor demographic mismatch in first-year law courses might substantively affect early-career earnings.

Appendix Table B3 estimates the fully-specified ordered-logit model separately by sex and race. The key results here are that (i) the gender pay gap exists for white, black, and Hispanic lawyers and (ii) that the wage-GPA gradient exists in the male, female, white, and black subsamples. Appendix Table B4 similarly shows that the wage-GPA gradient exists for graduates of all law schools outside the US News Top-10. This is consistent with results reported in Oyer and Schaefer (2016). The US News rank of the law school studied in the current paper falls in the 21-100 range (column 3), for whom grades are quite important.

[17] The AJD is a representative survey of new law-school graduates, conducted by the American Bar Foundation, in 2002, 2007, and 2010. See `http://www.americanbarfoundation.org/publications/afterthejd.html` for further information.

Table B1: Descriptive Ordered-Logit Earnings Regressions: Coefficient Estimates

	(1)	(2)	(3)	(4)
Female	-0.45***		-0.49***	-0.49***
	(0.07)		(0.07)	(0.07)
Black	-0.02		0.11	0.02
	(0.11)		(0.12)	(0.11)
Hispanic	-0.10		0.04	-0.01
	(0.11)		(0.12)	(0.12)
Asian	0.62***		0.63***	0.42***
	(0.13)		(0.14)	(0.13)
Other Race	0.06		0.10	0.06
	(0.19)		(0.19)	(0.20)
> 3.75 GPA		1.82***	1.90***	1.68***
		(0.16)	(0.16)	(0.18)
3.5 - 3.74 GPA		1.59***	1.65***	1.42***
		(0.13)	(0.13)	(0.14)
3.25 - 3.49 GPA		1.04***	1.09***	0.87***
		(0.13)	(0.13)	(0.13)
3.0 - 3.24 GPA		0.56***	0.59***	0.48***
		(0.12)	(0.12)	(0.12)
Missing GPA		1.34***	1.35***	1.16***
		(0.10)	(0.10)	(0.10)
Top 10 Law School				2.21***
				(0.15)
11-20 Law School				1.45***
				(0.13)
21-100 Law School				0.38***
				(0.07)
Observations	3785	3892	3785	3755

Note: Each column represents a different model specification. Cut points not shown. Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Descriptive Ordered-Logit Earnings Regressions: Average Partial Effects

	(1) 0-39K	(2) 40-49K	(3) 50-59K	(4) 60-74K	(5) 75-99K	(6) 100-124K	(7) 125-149K	(8) >150K)
Female	0.04***	0.04***	0.02***	0.00**	-0.02***	-0.02***	-0.02***	-0.04***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Black	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Hispanic	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Asian	-0.03***	-0.04***	-0.02***	-0.00*	0.02***	0.02***	0.02***	0.03***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Other Race	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00
	(0.02)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
> 3.75 GPA	-0.13***	-0.15***	-0.07***	-0.01**	0.06***	0.07***	0.08***	0.13***
	(0.02)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
3.5 - 3.74 GPA	-0.11***	-0.13***	-0.06***	-0.01**	0.05***	0.06***	0.07***	0.11***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
3.25 - 3.49 GPA	-0.07***	-0.08***	-0.03***	-0.00**	0.03***	0.04***	0.04***	0.07***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
3.0 - 3.24 GPA	-0.04***	-0.04***	-0.02***	-0.00*	0.02***	0.02***	0.02***	0.04***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Missing GPA	-0.09***	-0.10***	-0.05***	-0.01**	0.04***	0.05***	0.06***	0.09***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Top 10 Law School	-0.17***	-0.20***	-0.09***	-0.01**	0.08***	0.09***	0.11***	0.18***
	(0.01)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
11-20 Law School	-0.11***	-0.13***	-0.06***	-0.01**	0.06***	0.06***	0.07***	0.12***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
21-100 Law School	-0.03***	-0.03***	-0.02***	-0.00**	0.01***	0.02***	0.02***	0.03***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Observations	3755	3755	3755	3755	3755	3755	3755	3755

Note: Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Descriptive Ordered-Logit Earnings Regressions by Demographic Background: Coefficient Estimates

	(1) Male	(2) Female	(3) White	(4) Black	(5) Hispanic	(6) Asian
Female			-0.49***	-0.97***	-0.65**	-0.28
			(0.08)	(0.23)	(0.27)	(0.25)
Black	0.15	-0.09				
	(0.16)	(0.16)				
Hispanic	0.04	-0.09				
	(0.17)	(0.18)				
Asian	0.27	0.53***				
	(0.20)	(0.18)				
Other Race	0.13	-0.03				
	(0.28)	(0.29)				
> 3.75 GPA	1.62***	1.70***	1.69***	17.99***	0.61	1.44**
	(0.29)	(0.22)	(0.19)	(1.07)	(0.84)	(0.59)
3.5 - 3.74 GPA	1.54***	1.31***	1.48***	3.34***	2.86***	0.62
	(0.19)	(0.20)	(0.16)	(0.68)	(0.81)	(0.56)
3.25 - 3.49 GPA	0.92***	0.83***	0.87***	2.41***	0.40	0.48
	(0.18)	(0.19)	(0.15)	(0.64)	(0.53)	(0.35)
3.0 - 3.24 GPA	0.43**	0.53***	0.51***	0.83**	0.42	-0.10
	(0.17)	(0.17)	(0.14)	(0.39)	(0.43)	(0.41)
Missing GPA	1.15***	1.18***	1.27***	1.00***	0.71**	
	(0.14)	(0.15)	(0.12)	(0.29)	(0.33)	
Top 10 Law School	2.26***	2.15***	2.15***	2.49***	3.02***	2.08***
	(0.20)	(0.24)	(0.18)	(0.41)	(0.50)	(0.45)
11-20 Law School	1.47***	1.43***	1.51***	1.53***	0.85	1.03***
	(0.17)	(0.20)	(0.16)	(0.38)	(0.54)	(0.34)
21-100 Law School	0.34***	0.41***	0.42***	0.36	0.33	0.13
	(0.10)	(0.11)	(0.08)	(0.28)	(0.30)	(0.30)
Observations	1995	1760	2703	330	312	341

Note: Each column represents a different model specification. Cut points not shown. Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B4: Descriptive Ordered-Logit Earnings Regressions by Law School Rank: Coefficient Estimates

	(1) Top 10	(2) 11-20	(3) 21-100	(4) Outside 100
Female	-0.55**	-0.44**	-0.44***	-0.60***
	(0.24)	(0.20)	(0.10)	(0.12)
Black	0.07	0.11	-0.07	0.06
	(0.33)	(0.28)	(0.17)	(0.23)
Hispanic	0.42	-0.60	-0.00	0.10
	(0.37)	(0.37)	(0.17)	(0.27)
Asian	0.57	0.23	0.33*	0.72***
	(0.37)	(0.30)	(0.20)	(0.27)
Other Race	1.13*	-0.36	-0.28	0.67
	(0.64)	(0.46)	(0.26)	(0.41)
> 3.75 GPA	-0.12	2.72***	1.98***	1.54***
	(1.02)	(0.43)	(0.26)	(0.31)
3.5 - 3.74 GPA	0.80	2.15***	1.62***	1.33***
	(0.85)	(0.38)	(0.21)	(0.26)
3.25 - 3.49 GPA	0.97	1.50***	1.07***	0.50**
	(0.82)	(0.42)	(0.20)	(0.21)
3.0 - 3.24 GPA	0.42	1.99***	0.42**	0.41**
	(0.81)	(0.41)	(0.19)	(0.18)
Missing GPA	0.52	2.40***	1.33***	0.90***
	(0.78)	(0.33)	(0.16)	(0.16)
Observations	370	467	1737	1181

Note: Each column represents a different model specification. Cut points not shown. Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.