The impact of instructor rank on required first-year law course outcomes

Seth Gershenson[a,b]

Erdal Tekin[a,b,c]

Raymond Zuniga[d]

Abstract

A striking development in the U.S. higher education system in the last several decades is the changing composition course instructors, highlighted by a reduced reliance on tenure-track or tenured faculty and increased use of "teaching faculty" who generally have higher teaching loads, lower pay and research expectations, less job security, and no clear route to tenure. An open question in legal education is whether faculty of different rank are equally effective in the classroom. We provide novel insights into the relative effectiveness of instructors of different rank in a top-100 law school. Specifically, we compare the impacts of tenured, tenure-track, and "teaching" faculty on student-by-course outcomes in required first-year courses. Our empirical analysis employs student-by-course level administrative data from a top law school. We obtain credibly causal estimates by exploiting conditionally random variation in first-year classroom assignments complemented by a student-fixed effects research design. Students are about 4 percentage points (12%) more likely to receive good grades (As) from tenure-track faculty than from instructors of other ranks. Interestingly, this phenomenon is driven almost exclusively by the course grades of male students, particularly white male students. And while non-tenure line and tenured faculty exhibit similar grading patterns, non-tenure line faculty seem to inspire more interest in the subject matter, as their students are more likely to take future courses in the subject area and less likely to drop the course than are students taught by tenured and tenure-line faculty. While these results are likely causal in the sense that they are not due to sorting of students into courses, it is not clear whether these effects are driven by variation in teaching quality, implicit or other biases on the part of instructors than affect student engagement, or general differences in grading practices.

Notes: (a) American University, (b) IZA, (c) NBER, (d) Virginia Tech University. Corresponding author contact information: Gershenson gershens@american.edu. Tekin: tekin@american.edu. Zuniga: raymondz@vt.edu. The authors are thankful for financial support from the Access Group Research Grant Program. Opinions reflect those of the authors and not necessarily those of the granting agency.

1. Introduction

Over the last several decades, the composition of the academic workforce in U.S. higher education has undergone a significant shift, highlighted by a steadily increasing reliance upon non-tenure-track faculty teaching courses. Such positions generally have lower expectations for conducting scholarly research, higher teaching loads, and lower salaries. Some come with a one, two, or even five-year commitment, but others may operate on a semester-to-semester or even course-to-course basis. The main factors driving this development are desires to reduce expenses and/or increase flexibility, as there are often sharp disparities in salary and course-load between non-tenure track instructors and their tenure-track or tenured counterparts.

The shift away from employing tenure-line faculty in the classroom is notable in all types of postsecondary institutions. For example, between 2003 and 2013, the share of "teaching faculty" (i.e., non-tenure line) faculty members increased by 2 percentage points in private research-intensive institutions and by 17 percentage points in public four-year colleges and universities (Hurlburt and McGarrah 2016). Tenured and tenure-track faculty now account for less than 1/3 (29 percent) of the academic workforce in higher education (Yakoboski 2015).

As the reliance on non-tenure line faculty has grown, so to have concerns about the possible consequences to student learning. Critics argue that since non-tenure line instructors are typically less attached to campus life, less likely to hold a Ph.D. , and less engaged in scholarly research, faculty governance, and student advising, instructional quality in institutions that rely on them suffers and students will be worse off as a result (Haeger 1998; MLA 2003). Another concern is that contingent faculty members' concerns about having their contracts renewed might lead them to engage in practices that satisfy students but reduce learning, such as decreasing academic rigor or inflating grades in their classes (Sonner, 2000, Schuman, 2014).

Of course, the same concerns about course rigor and grading standards apply to tenure-line faculty too, whose merit reviews and tenure and promotion decisions are at least partly influenced by student assessments of their teaching. More importantly, there are also reasons to believe that non-tenure line faculty who have actual work experience outside of the academy can bring fresh perspectives, job and internship opportunities, and real life experiences into the classroom in ways that career academics cannot (Kirstein, 2015; Leslie and Gappa, 1995). This is all to say that whether tenure-line faculty are more or less effective in the classroom than their non-tenure line counterparts is theoretically ambiguous and is thus an empirical question.

However, while much has been written about the potential consequences of non-tenure line faculty for learning and student success, there is relatively little credible empirical evidence on the matter, particularly in professional school settings, both because of data limitations and the fact that obtaining credibly causal estimates is difficult given the selection problems inherent in higher education: students generally choose their classes just as department heads and deans generally choose which faculty teach which sections. A handful of exceptions, which we describe in more detail below, include the studies by Ehrenberg (2013), Figlio et al. (2015), Bettinger and Long (2010), Xiaotao and Xu (2019), Xu (2019), Xu and Solanki (2019). Still, as non-tenure line faculty have become the norm in U.S. colleges and universities, there is little sign of an end to debates about the relative merits and drawbacks of this strategy. Therefore, the need remains for credible evidence on the impact of "teaching faculty" on student outcomes.

In this paper, we extend the literature on the importance of faculty rank in the classroom by providing credible evidence on the causal impacts of tenured, tenure-track, and teaching faculty on course- outcomes in first-year required law school classes. Focusing on law schools is especially important because law schools might wish to further increase the usage of non-tenure

line instructors manage costs (Ehrenberg 2013). As job prospects and salaries for newly minted lawyers have levelled off in recent years, the lower costs of non-tenure line faculty offer law schools an attractive way to reduce expenditures to offset enrollment declines, increase financial aid offers, or reduce tuition. In fact, there are proposed rule changes by the American Bar Association (ABA) for law school accreditation that would may make it easier for law schools to hire and staff more classes with contingent faculty. For example, the ABA is considering eliminating a requirement that full-time faculty teach at least half of every law school's upper-level courses (Fain 2017). Therefore, there is even more urgency for credible evidence on the relative performance of various ranks of faculty to better inform law schools about the consequences such decisions.

Furthermore, determining whether non-tenure line instructors affect students' academic performance, particularly among students from backgrounds that are underrepresented in the legal profession, will improve our understanding of the barriers to obtaining law degrees and employment in the legal profession faced by such students even after they have been admitted to law school. Identifying the educational inputs under law schools' control that influence the performance and decisions of female and racial and ethnic minority law students is important, as even among similarly credentialed students at top law schools, women and minorities often feel alienated in law school classrooms and are less likely to graduate than white men (e.g., Banks, 1988; Guinier et al., 1994). This is further motivated by recent research showing that assignment to same-race or same-gender instructors significantly improves students' performance in first-year required law courses (Birdsall et al. 2020). These sorts of disparities in law school experiences likely contribute to the underrepresentation of females and racial and ethnic

minorities in the legal profession, particularly in the highest paying, most visible positions (Holder, 2001).

To investigate the relative efficacy of law school faculty of different ranks in facilitating law school student success, we utilize rich student-by-course administrative data that are linked to instructor data from a top-100 law school (LS) in the U.S. for the period between 2001-02 and 2012-13. A key feature of the LS we study is that for most of this time period, first-year law students were assigned to specific section groups for their required courses in a quasi-random fashion. As a result, individual students had no freedom to choose course sections, and therefore could not choose instructors of a certain rank. This conditionally random assignment allows us to circumvent the usual selection problem inherent in higher education settings.

2. Literature Review

To date, only a few credibly causal analyses of the impact of instructor rank have been conducted, all in the undergraduate context. First, Bettinger and Long (2010) address this question using data on all full-time, first-time, first-year college students in a sample of 12 public postsecondary institutions in Ohio. The authors use a quasi-experimental instrumental-variables strategy that exploits variation over time in the faculty composition of specific departments. They find that contingent faculty perform no worse than their tenured (and tenure-track) counterparts in the classroom. If anything, the authors find that contingent faculty are slightly more effective when it comes to promoting subsequent course enrollments.

Second, Figlio et al. (2015) address this question using detailed administrative data from eight cohorts of incoming first-year undergraduates at Northwestern University. The authors use a quasi-experimental two-way fixed effects strategy that makes both within-student and within-

second semester class comparisons. The authors find that contingent faculty are more effective than their tenured (and tenure-track) counterparts in terms of shaping students' subsequent course taking *and* in improving students' performance in subsequent courses in the same field. Despite using a different methodology and using data from a more selective university, these results are broadly consistent with those reported by Bettinger and Long (2010).

Finally, a more recent set of studies by Di Xu and colleagues (Xiaotao & Xu 2019; Xu 2019; Xu & Solanki 2019) yields a more mixed set of findings, though two of these studies focus on two-year community colleges, where the student-body and faculty do not resemble those at more selective universities like Northwestern. Specifically, Xu (2019) studies community colleges in a state system and finds that students in classrooms headed by part-time faculty tend to receive higher grades but then perform worse in subsequent, related courses. Xiaotao and Xu (2019) find similar results in a broader sample that includes public four-year institutions as well. However, the authors go on to differentiate between non-tenure line faculty who are on temporary versus longer-term contracts, which is potentially important given that critics of increasing non-tenure line faculty often focus on their relative lack of attachment to the university. Indeed, the authors find that the negative effects on subsequent outcomes are stronger for non-tenure line faculty on temporary contracts.

Motivated by the idea that job security might improve the teaching quality of non-tenure line faculty, some schools have begun to create a separate tenure track for teaching-only positions that combine the security of tenure with the higher-course loads and lower pay associated with non-tenure line positions. Xu and Solanski (2019) study such a program at one University of California campus. They generally find no differences by faculty rank in student outcomes.

Of course, these results do not necessarily generalize to professional- or law-school contexts, especially since results seem to vary between two- and four-year colleges and universities. Moreover, law students tend to be older, more mature, and have a clearer sense of professional direction. The educational environments are different too, as class sizes and teaching styles in first-year law school courses differ from those in first-year undergraduate courses. We contribute to the extant literature by studying the role of faculty rank in a top-100 law school, where plausibly causal estimates are obtained due to quasi-random first-year class assignments.

To our knowledge, the only study to broach this topic in the law school context is Ehrenberg (2013), who provides suggestive evidence that at the institution level, differences in first-year attrition rates and in bar passage rates are not attributable to the share of "full time" versus "part time" faculty. However, the author clearly articulates the myriad limitations of this initial analysis: the data did not distinguish tenured/tenure-line faculty from full-time non-tenure line teaching faculty nor did the data identify the share of full-time faculty teaching in crucial first-year courses. Another limitation of addressing this important question at the institution level regards within-institution sorting of students to classrooms, which might bias this sort of descriptive analysis. The present paper addresses some of these shortcomings by distinguishing between tenure-track, tenured, and non-tenure line faculty and by leveraging quasi-random class assignments to remove the threat of selection bias.

3. Data and Institutional Details

To investigate the relative efficacy of law school faculty of different ranks in facilitating law school student success, we utilize rich student-by-course administrative data from a private,

top-100 law school (LS). The LS enrolls approximately 1,000 students per year, on average, and employs approximately 200 full- and part-time faculty. While this study focuses on only one law school, the LS is representative of the type of selective (top 100) but not elite (top 10) law school attended by many students in the U.S. The LS is also relatively diverse in terms of student demographics, which provides power to conduct heterogeneity analyses.

We focus on required first-year courses for three reasons. First, entering students take the same set of courses during their first two semesters of law school, except for one elective in the spring semester. Second, first-year courses are assessed using an anonymous grading system, which allows for faculty adjustment of grades only at the last step, and grades are the main outcome that we analyze. Finally, at least in some years, student assignments to specific class-sections, made by LS advisors and administrators, were quasi-random. We further exclude a required legal writing class from the analytic sample because the way it is taught involves multiple instructors.

About 3 to 6 sections of the remaining courses are offered in semesters the course is offered, which leads to some within-course variation in instructor rank in that some full time teaching faculty may teach core doctrinal first year courses . We verify, and exploit, the conditionally random variation in rank that results from the first-year assignment process in the empirical analysis. These required first-year courses include subjects such as civil procedure, constitutional law, contract law, property law, torts.

The administrative data provide course titles and section numbers, instructors assigned to each course section, and final grades earned by students, which we use to examine course-specific completion and performance (i.e., grade) outcomes. We further leverage this course-specific information, along with other details available in the administrative data, to examine

future (year two) course-taking in the same legal concentration (i.e., constitutional, criminal, litigation, and property). Other valuable information included in the administrative data allows us to test for heterogeneous effects by student demographic characteristics, control for student-level covariates that may have been part of the conditionally random assignment process, and perform a balance test using data on student background to further verify the conditionally random course assignment process. We also observe, crucially for this analysis, instructor rank and experience at the LS by academic year. This changes over time, of course, as tenure-track faculty are awarded tenure and occasionally non-tenure line faculty transition to the tenure track.

3.1 Sample and Summary Statistics

Our goal is to estimate the effect of instructor rank in required first-year courses on course-specific outcomes. Our primary analytic sample includes a little over 30,000 student-course observations of first-year, full-time law students entering the LS between academic years 2001-02 to 2012-13.[1] Table 1 summarizes the student-course data for the full sample and separately by instructor rank. Because first-year courses must be taught by full-time faculty (and not contingent adjuncts), the non-tenure line category includes full-time teaching faculty who are generally on multi-year appointments. The majority of student-course sections were taught by tenured faculty (77%), with non-tenure line (17%) and tenure-track faculty (6%) rounding out the rest. The top panel summarizes student socioeconomic and demographic backgrounds. Columns 2-4 show how student backgrounds varied across instructor type. The general balance in student background across columns 2-4 supports the contention that class assignments were approximately random.

[1] All sample sizes are rounded to nearest 100 to preserve anonymity.

The middle panel of Table 1 summarizes course-specific student outcomes and how these varied by instructor rank. About 1/3 of students in a given course receive an A or A-, while only 6% receive lower than a B-. These numbers vary across instructor rank, however, in a way that foreshadows the main empirical result: tenure-track faculty award more good grades and fewer bad grades than their colleagues. Their students are also less likely to take a future elective course in the subject and more likely to drop the course than the students of other types of instructors. The future course outcome regards taking a subsequent elective in a related subject. So while the instructors we study are teaching in first-year required sections, the outcome variable in future-course models regards enrollment in a subsequent elective. Course drops are quite rare because these are required courses and dropping or otherwise changing sections is prohibited, unless there is a documented emergency or the student resigns from the LS. While suggestive, we will verify whether these differences can be interpreted as casual effects using the econometric techniques described in the next section.

Table 2 summarizes information describing the nearly 500 unique course sections in the main analytic sample overall, and by instructor rank. The average class size was about 80, and this is relatively balanced across instructor rank. So is the classroom's demographic composition, again suggesting that classroom assignments were approximately random. The eight courses that were required at some point during the sample period are less evenly represented across faculty ranks. These differences are due to changes in faculty availability from semester to semester and personnel specialties and preferences, which are decidedly not randomly distributed by rank, so all models will condition on course, or even course-by-term fixed effects, as explained in the methods section below.

Table 3 summarizes the unique instructors observed in the analytic sample. Unfortunately, missing data is an issue here, even for variables like experience that are supplied in the administrative data. This is particularly an issue for non-tenure line instructors. The models account for missing data by including a "missing" category. It is worth noting, however, that aggregated information found on the LS's website corroborates the basic descriptives presented here. Looking across the columns of Table 3, there are some important differences in instructor background by instructor rank. The source of some is intuitive, as tenure-track faculty have less experience than tenured faculty. There are also demographic differences, with more female and Latino tenure-track than tenured faculty, and no Black non-tenure line faculty. Once again, this highlights the importance of adjusting for instructor demographics in the empirics.

4. Research Design

Any attempt to obtain a credible estimate of the causal impact of faculty rank on student outcomes needs to address the potential endogeneity of student-instructor matches. Specifically, students may self-select into courses taught by a particular faculty type, and to the extent that the factors determining the choices of students are correlated with their performance, comparisons of student outcomes taught by various faculty types may reflect not only the performance of these faculty, but also the differences in student attributes. For example, some students may prefer to take their courses from tenured or tenure-track faculty on the assumption that these faculty are more effective teachers. Alternatively, the courses taught by non-tenure line faculty may be offered at particular times that may appear to certain types of students (Bettinger and Long, 2010).

We situate our analysis in the required first-year courses of LS, as the quasi-random assignment process ensures the internal validity of our results. This potentially comes at the expense of the external validity of our findings, due to reliance on a single institution, though we argue that even though LS is unique in a few ways, it is still largely comparable to other selective law schools in the top 100 that many students attend.

First year students in the LS are clustered into different sections and students in each section generally take the required first-year courses as a group. The formation of these clusters is conditionally random, in the sense that in some years some effort was made to create demographically balanced clusters. This means that conditional on some easily observed pre-admissions criteria, assignments to faculty of different rank should be as good as randomly assigned. Moreover, that the process was conducted by LS administrators suggests that students themselves could not sort into courses that are taught by a particular type of faculty, at least in their first year. See Birdsall et al. (2020) for further verification that the first-year assignments were as good as randomly made.

We proceed by modeling student-course outcomes (y) as

$$y_{ics} = \varphi_i + \omega_c + \delta RANK_s + \beta X_{isc} + \varepsilon_{isc}, \quad (1)$$

where i, c, and s index students, courses, and sections; φ is an unobserved student effect; ω is a set of eight course indicators, or fixed effects (FE), for each of the eight required first-year courses included in the analytic sample; RANK includes indicators for whether the section instructor was non-tenure line or tenure track, with tenured instructors being the omitted reference group; X is a vector of student, instructor, and course characteristics including students' cohort of entry, age, race, sex, home zip code characteristics, and LSAT score, the class size, and instructor characteristics such as experience at the LS, and demographics; and ε is an

idiosyncratic error term. The student effect φ can be ignored, treated as a random effect, or treated as a random effect. When treated as a FE, this subsumes the student controls, which do not vary across sections. We cluster standard errors at the section level because the treatment of interest (instructor rank) varies at the section level; however, the main results are robust to clustering the standard errors along other dimensions such as by student or instructor.

The parameters of interest are the elements of the vector δ, which measure the relative impacts of faculty rank on student outcomes. If the student assignments were made conditional only on some basic observed student characteristics, then student FE are unnecessary. Specifically, the intuition of the Hausman Test (Hausman 1978) suggests that if assignment were conditionally random, then estimates of δ should be robust to whether the student effect φ is modeled as a fixed or random effect, and we will estimate the model both ways to verify this.

5. Results

5.1 Main Results

Table 4 reports estimates of the baseline model specified in equation (1), which treats the unobserved student effect as a fixed effect and omits tenured faculty as the reference group, for five course-specific outcomes. Column 1 shows no significant effects of faculty rank on course grades, as measured on the four-point grading scale. However, the lack of a precisely estimated average effect on the numeric course grade could be misleading, as this is a not a continuous variable and the small but positive point estimate on tenure-track could be due to the effect of faculty rank 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 "poor" grades, respectively. Column 2 shows that

having a tenure-track instructor increased the chances of receiving an A or A- by about 4 percentage points, or 12%, relative to having either a tenured or non-tenure line instructor.[2] Interestingly, there is no statistically or practically significant difference between tenured and non-tenure line faculty. Nor, as column 3 shows, does faculty rank significantly affect the chances of receiving a poor grade (lower than B–).

Column 4 looks at the effect of faculty rank on student interest in the topic, as measured by whether the student takes at least one subsequent elective course in the same subject. Relative to tenured faculty, non-tenure line faculty significantly increase the likelihood that students take another course in the subject by about 2.5 percentage points, or 3%. Students of tenure-track instructors, on the other hand, are almost 3 percentage points less likely to take a subsequent course in the topic area, though this effect is less precisely estimated.[3]

Finally, column 5 of Table 4 considers dropping the course as an outcome. We see significant differences in the likelihood of dropping the course by instructor rank, with drops more likely for tenure-track instructors and less likely for non-tenure line instructors. The point estimates are small absolute terms, but this is because of the rarity of course drops; these effects would nearly double the course drop rate.[4]

Together, the results in table 4 indicate that faculty rank does influence law students' course-specific outcomes in required first-year courses. Relative to students of tenured faculty, students of tenure-track faculty are more likely to receive good grades but also show less interest

[2] This result is robust to instead defining "good grade" as receiving an A in the class.
[3] Note that this model is only estimated for the subsample of courses in which there is a potential elective in the same subject. The finding in column 2 of Table 4 that tenure-track faculty assign good grades more often than tenured faculty holds in this subsample as well.
[4] Course drops are rare because these are required courses and there are relatively rigid first-year requirements for progressing in the program.

in the subject matter: they are less likely to take future courses in the subject and are more likely to drop the course. Meanwhile, students of tenure-track faculty receive the same grades, on average, as students of tenured faculty, but they show significantly more interest in the topic than do the students of either tenured or tenure-track faculty. We discuss why this may be in the conclusion.

5.2 Robustness Checks

In addition to indicating that students' decisions are affected instructor rank, the course-dropping results raise a concern about how to interpret the effects of instructor rank on course grades. The reason is that course grades are not observed for students who dropped the course, which creates non-random selection into the course-grades analytic sample. Accordingly, in the spirit of Lee (2009), we construct bounds for the estimated effects of faculty rank on course grades and verify that the results reported in Table 4 are not driven by this nonrandom sample selection.

The more obvious threat to validity, discussed in sections 3 and 4, is whether the baseline model presented in equation (1) adequately controls for selection into courses. Table 5 reports estimates of several variants of equation (1), for the likelihood of receiving a good grade, that progressively control for richer dimensions of student background. The idea here is that once the basic set of conditioning variables used by administrators to assign students to courses is included is included in the model, additional controls (e.g., student FE) should not matter in the sense that the estimated effect of faculty rank should be robust to their inclusion.

Column 1 reports estimates of a naïve version equation (1) that includes no controls whatsoever, other than the eight course fixed effects and cohort indicators. This leads to an

estimated coefficient on tenure-track that is twice as large as that generated by our baseline model (reported in column 2 of table 4) and a positive, significant effect of non-tenure line faculty. But we know this model is likely biased, as at least in some years observed student characteristics were used to shape course assignments. Accordingly, the model estimated in column 2 adds the full set of student, section, and instructor controls to the model, but still leaves the unobserved student effect in the error term. Adding these controls brings the estimated coefficients back in line with the baseline (student FE) estimates from table 4. This suggests that the student FE is unnecessary and implies that sorting only occurred on observed, and not on unobserved, student characteristics.

In columns 3 and 4 of table 5 we formalize this idea by implementing a Hausman test of the RE versus FE estimators. The RE estimates are nearly identical to the OLS estimates and are more precisely estimated. The baseline FE estimate on tenure-track, reproduced in column 4 of Table 5, is 0.005 smaller than the RE estimate, but this difference is not statistically significantly. This reaffirms that the student FE are not necessary, that there was no selection into courses on unobserved student characteristics, and that the baseline estimates can be given a causal interpretation. Moreover, the baseline student-FE estimates are conservative in the sense that they are marginally smaller and less precisely estimated than the RE estimates.

Finally, columns 5 and 6 of table 5 replace the course FE with course-by-term FE. The idea here is that there may be semesters when there is limited or even no variation in instructor rank in a given course. We account for this by re-estimating the baseline model on the sample of course-terms with adequate variation, and again find that the effect remains about 0.04.

5.3 Heterogeneity

To this point, we have established that, on average, tenure-track faculty assign more good grades and that this is likely a causal effect of faculty rank as opposed to tenure-track faculty having better students or teaching easier subjects. We have seen on difference in the assignment of good grades between tenured and non-tenure line faculty. Given the general diversity of the student body and how students from underrepresented backgrounds often experience law school differently than their white and male counterparts (Birdsall et al. 2020), we now repeat the main analysis separately by student demographics, to see whether the pooled analysis masked any substantial heterogeneity by student background. These estimates are reported in Table 6.

Columns 1 and 2 of Table 6 estimate the baseline model separately by student gender. Here, we see that the effect of tenure-track faculty on good grades was entirely driven by male students receiving more good grades from tenure-track faculty. The effect of non-tenure line faculty is also driven by male students, though the effect is imprecisely estimated. Columns 3 and 4 similarly estimate the baseline model separately for white and nonwhite students, respectively. Once again, we see that the tenure-track effect on good grades is almost entirely driven a historically advantaged group: white students.

Columns 5 through 8 of table 6 further divide the sample by sex and race. Here, we see that the tenure-track effect applies to all males, and while the effect is slightly large for white men than for nonwhite men, this difference is not statistically significant. For women of any race, however, there is no effect of faculty rank.

6. Discussion

The past several decades have seen increases in the reliance of non-tenure line "teaching faculty" in higher education, including in professional schools. This is largely predicated by the lower costs and increased flexibility of non-tenure line faculty. However, this comes at the cost of potentially less engagement with the university, school, and students which may harm the student experience as a result. Whether instructors of different ranks are differentially effective in the classroom is an open question in legal education. We provide novel insights into this question using rich administrative data from a top-100 law school.

When comparing the impacts of tenured, tenure-track, and "teaching" faculty on student-by-course outcomes in required first-year courses, we see that students are about 4 percentage points (12%) more likely to receive good grades (As) from tenure-track faculty than from instructors of other ranks. Interestingly, this phenomenon is driven almost exclusively by the course grades of male students, particularly white male students. And while non-tenure line and tenured faculty exhibit similar grading patterns, non-tenure line faculty seem to inspire more interest in the subject matter, as their students are more likely to take future courses in the subject area and less likely to drop the course than are students taught by tenured and tenure-line faculty.

These results are arguably causal in nature, as we exploit conditionally random variation in first-year classroom assignments. Moreover, we verify the conditional randomness of these assignments by reporting balance tests and by showing that the main results are robust to conditioning on student fixed effects. However, while these results are likely causal in the sense that they are not biased by student sorting into courses, the mechanisms through which these effects operate is unclear. This merits future research. For example, it is not clear whether these

effects are driven by variation in teaching quality, implicit or other biases on the part of instructors than affect student engagement, or general differences in grading practices.

References

Banks, T. L. (1988). Gender bias in the classroom. Journal of Legal Education, 137-146.

Bettinger, E. P., & Long, B. T. (2010). Does cheaper mean better? The impact of using adjunct instructors on student outcomes. *The Review of Economics and Statistics*, *92*(3), 598-613.

Birdsall, C., Gershenson, S., & Zuniga, R. (2020). The Effects of Demographic Mismatch in an Elite Professional School Setting. Forthcoming, *Education Finance and Policy*.

Ehrenberg, R. G. (2013). American Law Schools in a Time of Transition. *Journal of Legal Education*, *63*(1), 98-112.

Fain, Paul. 2017. More Law School Courses for Adjuncts? https://www.insidehighered.com/quicktakes/2017/03/29/more-law-school-courses-adjuncts.

Figlio, D. N., Schapiro, M. O., & Soter, K. B. (2015). Are tenure track professors better teachers? *Review of Economics and Statistics*, *97*(4), 715-724.

Guinier, L., Fine, M., Balin, J., Bartow, A., & Stachel, D. L. (1994). Becoming gentlemen: Women's experiences at one Ivy League law school. University of Pennsylvania Law Review, 1-110.

Haeger, J. D. (1998). Part-time faculty, quality programs, and economic realities. In D. W. Leslie (Ed.), *The growing use of part-time faculty: Understanding causes and effects* (pp. 81–88). San Francisco: Jossey-Bass.

Holder Jr, E. H. (2001). Importance of Diversity in the Legal Profession, The. *Cardozo L. Rev.*, *23*, 2241.

Hurlburt, S. and McGarrah, M., 2016. The shifting academic workforce: Where are the contingent faculty. *Delta Cost Project at American Institutes for Research. Retrieved from https://www.air.org/sites/default/files/downloads/report/Shifting-Academic-Workforce-November-2016.pdf*.

Kirstein, Rita. 2015. Colleges' Reliance on Part-time Faculty: Who Wins? Who Loses? https://www.noodle.com/articles/the-rise-of-part-time-faculty-and-its-effect-on-higher-ed131

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, *76*(3), 1071-1102.

Leslie, David W., and Judith M. Gappa, "The Part-Time Faculty Advantage," Metropolitan Universities: An International Forum 6:2 (1995), 91–102.

Modern Language Association, "Ensuring the Quality of Undergraduate Programs in English and Foreign Languages: MLA Recommendation on Staffing" (2002), http://www.mla.org.

Schuman, Rebecca. 2014 Confessions of a Grade Inflator: Between the grubbing and the blubbering, grading fairly is just not worth the fight. Slate. http://www.slate.com/articles/life/education/2014/05/why_professors_inflate_grades_because_their_jobs_depend_on_it.html

Sonner, B.S., 2000. A is for "adjunct": Examining grade inflation in higher education. *Journal of Education for Business*, *76*(1), pp.5-8.

Xiaotao, F., & Di Xu, R. (2019). Does contractual form matter? The impact of different types of non-tenure-track faculty on college students' academic outcomes. *Journal of Human Resources*, *54*(4), 1081-1120.

Xu, D. (2019). Academic performance in community colleges: The influences of part-time and full-time instructors. *American Educational Research Journal*, *56*(2), 368-406.

Xu, D., & Solanki, S. (2019). Tenure-Track Appointment for Teaching-Oriented Faculty? The Impact of Teaching and Research Faculty on Student Outcomes. *Educational Evaluation and Policy Analysis*, 0162373719882706.

Yakoboski, P. J. 2015. The Career Experience of Academics in Adjunct Faculty Positions. Trends and Issues. TIAA-CREF Institute. https://www.tiaainstitute.org/sites/default/files/presentations/2017-02/adjunct_career_experience_full.pdf

Table 1. Student-by-course summary statistics, by faculty rank

	All	Non-Tenure line	Tenure-track	Tenured
	(1)	(2)	(3)	(4)
Student Background				
Age in first semester	25.45	25.45	25.42	25.45
	(2.56)	(2.64)	(2.58)	(2.54)
Non-white Student	0.36	0.37	0.37	0.36
Race data missing	0.025	0.026	0.028	0.025
Female Student	0.582	0.580	0.601	0.581
Median income in home zip	75,826	76,901	80,919	75,202
	(32,196)	(32,204)	(33,349)	(32,066)
%Adults w/ BA in home zip	38.31	38.65	34.26	39.19
	(20.14)	(20.14)	(17.20)	(20.31)
Home zip code missing	0.034	0.033	0.024	0.035
International Student	0.003	0.003	0.002	0.003
Course Outcomes				
Course Grade (0-4)	3.266	3.270	3.364	3.258
	(0.489)	(0.515)	(0.464)	(0.485)
Good grade (A/A-)	0.323	0.345	0.413	0.311
Bad grade (C-F)	0.064	0.073	0.034	0.064
Take future course in subject	0.798	0.757	0.681	0.820
Dropped course	0.002	0.002	0.005	0.002
Faculty Rank				
Non-Tenure-line	0.17	1	0	0
Tenure-track	0.06	0	1	0
Tenured	0.77	0	0	1
Observations	30,500	5,200	1,800	23,500

*Note*s. Standard deviations shown in parentheses for non-binary variables. There is no missing information for student age and sex. Sample size is smaller for the "take a future course" oucome because it is only computed for subjects with a future elective to take.

Table 2. Course-level summary statistics, by faculty rank

	All	Non-Tenure line	Tenure-track	Tenured
	(1)	(2)	(3)	(4)
Class Size	80.39	77.78	81.28	80.96
	(17.41)	(18.70)	(13.94)	(17.33)
Course %Non-white	36.51	38.04	38.34	35.99
	(9.544)	(11.38)	(11.38)	(9.048)
Course %Female	58.26	58.33	61.24	58.03
	(10.71)	(12.70)	(8.87)	(10.30)
Civil Procedure	0.146	0.081	0.000	0.175
Civil Procedure II	0.053	0.012	0.000	0.068
Constitutional Law	0.153	0.322	0.276	0.101
Contracts	0.146	0.161	0.000	0.155
Criminal Law	0.163	0.172	0.207	0.158
Property I	0.146	0.046	0.345	0.155
Property II	0.049	0.023	0.069	0.054
Torts	0.142	0.184	0.103	0.135
Non-Tenure-line	0.185	1	0	0
Tenure-track	0.062	0	1	0
Tenured	0.754	0	0	1
Observations	470	90	30	350

*Note*s. Standard deviations shown in parentheses for non-binary variables.

Table 3. Instructor summary statistics, by instructor rank

	All	Non-Tenure line	Tenure-track	Tenured
	(1)	(2)	(3)	(4)
Years at institution	9.00	4.59	1.58	11.31
	(10.23)	(2.13)	(2.22)	(11.15)
Years missing	0.239	0.690	0.001	0.001
Black	0.123	0.000	0.111	0.152
Latino	0.046	0.100	0.111	0.022
Asian	0.062	0.000	0.333	0.022
White	0.769	0.900	0.444	0.804
Race missing	0.226	0.655	0.000	0.000
Female	0.477	0.700	0.778	0.370
Gender missing	0.226	0.655	0.000	0.000
Observations	80	30	10	50

*Note*s. Standard deviations shown in parentheses.

Table 4. Effect of instructor rank on course-specific outcomes

Outcome:	Grade (0-4) (1)	Good Grade (2)	Poor Grade (3)	Future Course (4)	Drop Course (5)
Non-Tenure-line	-0.002	0.008	0.008	0.025**	-0.005***
	[0.024]	[0.021]	[0.010]]	[0.012]	[0.002]
Tenure-track	0.010	0.043**	0.008	-0.029	0.003*
	[0.022]	[0.019]	[0.011]	[0.018]	[0.002]
H_0: $\delta_1 = \delta_2$				**	***
Observations	30,300	30,300	30,300	19,800	30,400

Notes. Robust standard errors clustered by course sections shown in brackets. The reference group is courses taught by Tenured instructors. A good grade is an A- or better; a poor grade is a C or worse. All models condition on course dummies, student fixed effects, and course and instructor controls. Column (4) is only estimated for the subset of courses that have subsequent electives in the same area. * p<.10, ** p<.05, *** p<.01

Table 5. Effect of instructor rank on good grade (A/A-) in course

	(1)	(2)	(3)	(4)	(5)	(6)
Non-Tenure-line	0.026**	0.007	0.006	0.008	0.008	0.006
	[0.012]	[0.020]	[0.011]	[0.021]	[0.032]	[0.028]
Tenure-track	0.085***	0.048***	0.048***	0.043**	0.037	0.036
	[0.017]	[0.017]	[0.013]	[0.019]	[0.027]	[0.022]
H_0: $\delta_1 = \delta_2$	***	*	***			
Observations	30,400	30,200	30,200	30,300	30,200	30,300
Student X	No	Yes	Yes	No	Yes	No
Section X	No	Yes	Yes	Yes	Yes	Yes
Instructor X	No	Yes	Yes	Yes	Yes	Yes
Student RE	No	No	Yes	No	No	No
Student FE	No	No	No	Yes	No	Yes
Course-by-term FE	No	No	No	No	Yes	Yes

Notes. Robust standard errors clustered by course sections shown in brackets. The reference group is courses taught by Tenured instructors. A good grade is an A- or better. All models condition on course fixed effects (FE). The estimates in columns 2 through 6 are not significantly different from one another. * p<.10, ** p<.05, *** p<.01

Table 6. Effect of instructor rank on good grade (A/A-) by student demographics

Subgroup:	Female	Male	Nonwhite	White	Nonwhite Female	Nonwhite Male	White Female	White Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Non-Tenure-line	-0.004	0.024	-0.001	0.013	-0.019	0.033	0.006	0.021
	[0.025]	[0.023]	[0.023]	[0.025]	[0.027]	[0.030]	[0.030]	[0.027]
Tenure-track	0.013	0.084***	0.014	0.058***	-0.014	0.067**	0.029	0.093***
	[0.022]	[0.026]	[0.024]	[0.021]	[0.030]	[0.030]	[0.027]	[0.031]
H_0: $\delta_1 = \delta_2$		*						*
Observations	17,600	12,700	10,700	19,500	7,100	3,700	10,600	9,000

Notes. Robust standard errors clustered by course sections shown in brackets. The reference group is courses taught by Tenured instructors. A good grade is an A- or better. All models condition on course dummies, student fixed effects, and course and instructor controls.
* p<.10, ** p<.05, *** p<.01