

ARE TENURE TRACK PROFESSORS BETTER TEACHERS?

David N. Figlio, Morton O. Schapiro, and Kevin B. Soter*

Abstract—This study makes use of detailed student-level data from eight cohorts of first-year students at Northwestern University to investigate the relative effects of tenure track/tenured versus contingent faculty on student learning. We focus on classes taken during a student's first term at Northwestern and employ an identification strategy in which we control for both student-level fixed effects and next-class-taken fixed effects to measure the degree to which contingent faculty contribute more or less to lasting student learning than do other faculty. We find consistent evidence that students learn relatively more from contingent faculty in their first-term courses. This result is driven by the fact that the bottom quarter of tenure track/tenured faculty (as indicated by our measure of teaching effectiveness) has lower "value added" than their contingent counterparts. Differences between contingent and tenure track/tenured faculty are present across a wide variety of subject areas and are particularly pronounced for Northwestern's average and less-qualified students.

I. Introduction

THE role of tenure in American higher education has been reduced dramatically in recent decades. In 1975, 57% of all faculty (excluding graduate students) were in the tenure system; by 2011 that figure had been cut almost in half, to 29%.¹ Some observers predict that the share of tenure track/tenured faculty will bottom out at between 15% to 20%, with tenure being largely limited to the flagship public and private research universities and the wealthiest of the liberal arts colleges.²

There is evidence that this trend accelerated after January 1, 1994, when mandatory retirement for faculty was abolished by federal law. Ehrenberg (2012) reports that between 1995 and 2007, the share of part-time faculty rose at almost all institutional types, while among full-time faculty, the movement away from the tenure system has quickened. Especially notable is the rise of the full-time, contingent faculty member at Ph.D.-granting universities. Their representation within the entire group of full-time faculty went from 24% to 35% at public doctoral institutions and from 18% to 46% at private nonprofit doctoral institutions.

Received for publication September 3, 2013. Revision accepted for publication March 4, 2015. Editor: Asim Ijaz Khwaja.

* Figlio: Northwestern University and NBER; Schapiro: Northwestern University; Soter: Greatest Good.

We are grateful to the Northwestern University Registrar's Office, Office of Admissions, and Office of Human Resources for providing the data necessary to carry out this analysis, and to numerous colleagues, seminar participants, and the editor and referees for many helpful suggestions. Caitlin Ahearn, Rebecca Figlio, and Christine Mulhern provided exceptional research assistance. All opinions and errors are our own.

¹ The American Association of University Professors website presents its Contingent Faculty Index summarizing data from the IPEDS Fall Staff Survey.

² There has been ongoing speculation about this topic in the educational press. Wilson (2010) is a good example.

This trend has led some observers to lament the potential blow to academic freedom dealt by the decline of tenure and to focus on the often challenging employment conditions under which many contingent faculty work (see, e.g., June, 2012, and Wilson, 2010). Further, McPherson and Schapiro (1999) point to efficiency gains from tenure; they outline its positive role in influencing the distribution of authority within colleges and universities.

While those considerations certainly have relevance in evaluating the impact of the growing demise of the tenure system, there is an educational outcome that may be measured more directly: Do undergraduates taught by contingent faculty members learn as much as those taught by tenure track/tenured faculty?

There have been a number of attempts to answer this question. On a national level, Ehrenberg and Zhang (2005) present evidence that hiring more part-time and contingent faculty lowers institutional graduation rates. This result is bolstered by Bettinger and Long (2006), who find a similarly negative effect on aggregate levels of persistence when they focus specifically on part-time adjuncts. These types of results indicate that even if contingent faculty are more popular with students—perhaps because of classroom behaviors that maximize student evaluations but not student learning—they nonetheless might not be successful in improving students' longer-term prospects.³ To date, however, there exists little evidence on the effects of faculty tenure track status on genuine student learning.

The limited existing evidence on the relative performance of tenure track/tenured professors versus faculty outside the tenure system makes it difficult for college and university decision makers to determine the optimal staffing of their classrooms. This is particularly relevant for research universities, which face a multitasking problem of maximizing an objective function that includes both the production of cutting-edge research and the provision of outstanding undergraduate teaching. While the paper closest to this one in the literature, Bettinger and Long (2010), presents a novel approach to measuring the effects of tenure line versus other instruction, their analysis is largely centered on institutions whose principal purpose is teaching. Further, their creative identification strategy uses short-term vacan-

³ Carrell and West (2010) show that instructors who have better student evaluations tend to produce lower levels of "deep learning."

cies in departments, essentially analyzing the effects on learning of transitory adjunct faculty. Their analysis therefore may not speak to the effects of part-time faculty with longer-term contracts and certainly does not address the effects of full-time contingent faculty. In addition, while they find some evidence that contingent faculty induce student interest in a subject, as measured by the likelihood that students take additional courses in that subject, they are not able to study how students perform in subsequent classes, an ideal way to see whether instructional quality has a lasting impact. When one observes only student evaluations of their instructors or the likelihood that students take more classes in the subject, it is difficult to judge whether one type of instructor is genuinely better at education—that is, whether they produce more “deep learning” in the words of Carrell and West (2010)—or whether they are just more popular. Hoffmann and Oreopoulos (2009) evaluate teacher quality in a Canadian research university setting, but, like Bettinger and Long (2010), they observe only the likelihood that students take additional classes in the same subject rather than observe their academic performance in future classes.⁴ They find no evidence that contingent faculty are either better or worse at inspiring students to take more classes in their subjects. Carrell and West’s (2010) analysis of professor quality examines follow-on classes and has outstanding internal validity as it relies on the random assignment of students to classes, but it is also based at an institution (the U.S. Air Force Academy) where teaching rather than research is the dominant function. In addition, Carrell and West do not directly take on the question of whether contingent faculty make better or worse instructors.

We bring to bear the first evidence within the research university setting regarding the undergraduate learning effects from different faculty types where we can observe student performance in subsequent classes in the same subject. Specifically, we examine the initial classes taken by first-term freshmen in eight cohorts of undergraduates at Northwestern University, a midsized research university that is one of the twenty-six private universities among the sixty-two members of the Association of American Universities and that consistently ranks among the most selective undergraduate institutions in the United States. At Northwestern, contingent faculty members tend to have stable, long-term relationships with the university, and a substantial majority are full time. This allows us to study the effects of contingent faculty at a major research university where these faculty members function as designated teachers (both full time and part time) with long-term relationships to the university.

⁴ Hoffmann and Oreopoulos do study student performance in courses with standardized tests shared across individual class sections in the same term as their measure of instructor quality, but they do not perform a head-to-head comparison of lecturers versus tenure track/tenured faculty members (their comparison controls for a measure of instructor quality) and are not able to follow students into future classes to gauge “deep learning.”

Our identification strategy involves observing whether a student who takes, say, introductory economics with a contingent faculty member and introductory political science with a tenure track/tenured professor in his or her first term at Northwestern is (a) relatively more likely to take a second political science class than another economics class and (b), conditional on taking more classes in both subjects, more likely to perform better in the political science class than in the economics class.

The answers to these questions should shed light on one of the most important outcomes relating to the dramatic change in the professorate: its impact on student learning.

II. Data and Methods

We make use of data on all Northwestern University freshmen who entered between fall 2001 and fall 2008, a total of 15,662 students.⁵ Our principal model for estimating the relationship between the tenure track/tenured versus contingent status of a student’s instructor and that student’s level of learning in that subject is

$$G_{icst+1} = \alpha_i + \gamma_{cst+1} + \beta L_{ist} + \varepsilon_{ist},$$

where, for student i taking a first-term class in subject s at time t , L represents whether the class taken is taught by a contingent faculty member and G represents that student’s grade (on a four-point scale) the next time the student takes a class in subject s . The subscript c pertains to a specific instructor-class-term-year combination, so the inclusion of a fixed effect γ_{cst+1} means that we are comparing the relative performance in subsequent classes in subjects A and B of a student who took a class in subject A with a contingent faculty member and subject B with a tenure track/tenured professor during his or her first term at Northwestern, holding constant all of the specifics of the subsequent classes in subjects A and B. This means that we are obtaining our identification from subjects where some first-term freshmen take classes from a tenure line professor and other first-term freshmen take classes from a contingent faculty member. We also estimate linear probability models without the next-class fixed effect but with student fixed effects where the dependent variable is whether the student takes another class in subject s . We cluster standard errors at the instructor level to account for potential within-instructor error correlation.⁶

We obtained data from several offices at Northwestern University for the purposes of this analysis. The registrar’s office provided us with student transcript data, including student grades, subjects, and instructor information; the

⁵ We limit our analysis to students who entered Northwestern in fall 2008 or before to give students sufficient time to complete their studies. Ninety-eight percent of students who ultimately earn an undergraduate degree at Northwestern do so within five years.

⁶ In a prior iteration of this paper, we clustered our standard errors at the individual student level (Figlio, Schapiro, & Soter, 2013). The standard errors are larger in the case in which we cluster at the instructor level, but the fundamental message is the same as before.

TABLE 1.—DESCRIPTIVE STATISTICS, BY NUMBER OF CLASSES TAKEN WITH CONTINGENT FACULTY IN FALL QUARTER OF FRESHMAN YEAR

	Count	Highest Academic Preparation	Middle Academic Preparation	Lower Academic Preparation	Mean SAT Score	Undeclared at Entry	Took Another Class in Subject	Mean Grade in Next Class in Subject
Full sample	15,662	17%	57%	26%	1392	17%	74%	3.39
By number of contingent faculty classes								
0 contingent	3,144	17%	56%	27%	1391	13%	72%	3.24
1 contingent	5,978	17%	58%	25%	1395	17%	72%	3.32
2 contingent	4,019	17%	56%	27%	1395	20%	74%	3.41
3 or more contingent	1,925	17%	57%	26%	1392	14%	82%	3.62
Only contingent	596	14%	51%	36%	1362	19%	77%	3.46

Data include all students who enrolled in their first quarter at Northwestern University during the fall terms between 2001 and 2008.

office of admissions provided information about the student's initial intended major and academic qualifications; and individual academic departments, as well as the office of human resources, confirmed the tenure track/tenured versus contingent status of all instructors.⁷ Table 1 presents some descriptive statistics of the population of Northwestern students. We divide freshman students into groups based on academic preparation: 17% of entering freshmen have the highest preparation level, 57% have middle preparation levels, and 26% have relatively low levels of preparation. The average SAT score (or converted ACT score) for beginning freshmen was 1392, and 17% of entering freshmen had not indicated an intended major at their time of entry to Northwestern. In 74% of cases, students took another class in a subject that they took during their first term of freshman year, and when they took the subsequent class, they averaged a grade of 3.39 on a 4-point scale.

We limit our analysis to first-term freshman students because our identification assumption is that students select their first classes with limited knowledge about instructor quality or characteristics. We further condition on student fixed effects because we are concerned that students who take classes with one type of instructor versus another may be relatively strong or weak students. The majority of students take at least one course with a contingent faculty member and at least one course with a tenure line professor during their first term at Northwestern; 20.1% of students take classes only with tenure track/tenured professors, and 3.8% take classes only with contingent faculty members.

Rows 2 through 6 of table 1 break down descriptive statistics by the number of first-term classes taught by contingent faculty. The small number of students who take only classes taught by contingent faculty tend to be somewhat weaker than the other students; 36% come from the bottom ranks of students (as opposed to 26% for the other 96.2 percent of students), and their SAT scores average 1362 (as opposed to 1393 for the other students). But among the 96.2% of students who take at least one class from a tenure line faculty member, there is no apparent relationship between the division of

tenure track/tenured versus contingent classes and initial preparation. All four groups have 17% with highest academic preparation, 25% to 27% with relatively low academic preparation, and average SAT scores between 1391 and 1395.

Foreshadowing our results, the four groups differ substantially in terms of their outcomes. The probability that a student takes another class in the subject generally increases with the number of contingent faculty classes the student takes in his or her first term at Northwestern, as does the grade earned in the subsequent class. This latter pattern is especially remarkable given that contingent faculty appear to induce relatively marginal students, who might have been expected to perform worse in subsequent classes, to take those classes nonetheless. The bulk of this paper explores these relationships in a more systematic manner.

III. Estimated Effects of Contingent Faculty on Subsequent Performance

Table 2 presents our basic results. The unit of analysis is the student-class pair for first-term freshmen at Northwestern. To provide a basis for comparison, we report basic OLS results in the first row of the table and then successively add layers of fixed effects. The left-most columns of the table are for all classes taken by all students, while the second set of columns restricts the analysis to the 89.9% of classes taken outside a student's intended major.⁸

As can be seen in the first row of table 2, the simple relationships between contingent faculty status of the teacher of a class and a student's likelihood of taking another class and the grade obtained in that next class in a subject are positive and strongly statistically significant. However, because these relationships could reflect unmeasured student characteristics, we compare subjects taken by the same student and estimate student fixed-effects models. The estimated relationships remain reasonably large in magnitude: a contingent faculty member increases the likelihood that a student will take another class in the subject by 7.3 percentage points (9.3 percentage points when limited to classes outside the student's intended major) and increases the

⁷ We exclude graduate students and visiting professors who hold faculty appointments at other institutions from our analysis. Our results are fundamentally unchanged if we include these two groups, regardless of whether we assign them to the tenure track/tenured category or the contingent category of instructor.

⁸ We show separate results for classes outside the student's major to isolate the group of students for whom the choice to take additional classes in the subject is most plausibly affected by the quality of the first professor they encounter.

TABLE 2.—ESTIMATED EFFECTS OF HAVING A CONTINGENT FACULTY MEMBER ON SUBSEQUENT COURSE TAKING AND PERFORMANCE IN THE SUBJECT: LINEAR PROBABILITY MODELS

	All Classes		Classes Outside Intended Major (89.9% of Classes)	
	Probability of Taking Next Class in Subject	Grade in Next Class Taken in Subject	Probability of Taking Next Class in Subject	Grade in Next Class Taken in Subject
OLS regression	0.077*** (0.030)	0.185*** (0.050)	0.085*** (0.032)	0.218*** (0.051)
Regressions with student fixed effects				
All first-year fall classes	0.073*** (0.024)	0.120*** (0.037)	0.093*** (0.026)	0.159*** (0.041)
Students with no choice (35.2%)	0.108*** (0.034)	0.095*** (0.043)	0.139*** (0.038)	0.148*** (0.050)
Regressions with student fixed effects and next-class fixed effects				
All first-year fall classes	NA	0.060*** (0.008)	NA	0.079*** (0.009)
Controlling for home country, experience in years	NA	0.039*** (0.010)	NA	0.064*** (0.010)
Controlling for home country, experience in six bins	NA	0.042*** (0.010)	NA	0.066*** (0.010)

Data include all students who enrolled in their first quarter at Northwestern University during the fall terms between 2001 and 2008. Each cell represents a different model specification. Standard errors adjusted for clustering at the instructor level are reported in parentheses beneath coefficient estimates. Next-class information is recorded for the first time a student takes a second class in a given subject area. Intended majors are recorded by the office of admissions. Coefficients are statistically distinct at ***1%, **5%, *10%. Data come from 15,662 students taking 56,599 first-quarter classes.

grade earned in that subsequent class by slightly more than one-tenth of a grade point (with a somewhat greater impact for classes outside the intended major).⁹

We can further restrict our analysis to students with “no choice”—classes that are always taught either by tenure track/tenured faculty or by contingent faculty during the entire time period considered. In this restriction, we explicitly eliminate the possibility that a student is “shopping” across instructors teaching a certain class. The idea here is

⁹ Almost all classes taught by contingent track faculty at Northwestern are taught by those with a longer-term relationship with the university. When we exclude the temporary lecturers and adjuncts, the estimates barely change. Trimming the “one-off” lecturers and adjuncts, we find that for all students and all classes, a contingent faculty member is estimated to increase the likelihood that a student takes another class in the subject by 7.5 percentage points and increases the grade by 0.12 grade points. The results are similarly nearly identical for all other rows in the table.

Ten percent of students (19% of student-class pair observations) took multiple courses in the same subject in their first quarter, and in 2.3 percent of cases, a student took classes taught by both tenure track/tenured and contingent faculty members in the same subject that term. If we assign the next class to both tenure track/tenured and contingent faculty members in the same subject, as we do in the specifications reported in table 2, this has the effect of biasing our estimates toward 0. Indeed, if we limit our analysis to students who took only one course in the subject during their first quarter at Northwestern, we estimate that a contingent faculty member increases the likelihood that a student takes another class in the subject by 7.9 percentage points and increases the grade by 0.14 grade points. We continue to report the more conservative estimates in the paper.

The results are also robust when we limit our analysis to each of the specific colleges (there are six undergraduate colleges at Northwestern) where the classes were offered. In student fixed-effects regressions, the relationship between contingent faculty status and the probability that a student will take another class in the subject is positive and statistically significant in three of the four colleges (arts and sciences, music, and engineering, but not communications) that teach almost all of the first-term freshman students, and the relationship between contingent faculty status and the grade earned in the next class in the subject is positive and statistically significant in all four colleges. In section V of the paper, we also break down our results by the grading standards of the subjects and the qualifications of students who intend to major in those subjects.

that it is conceivable that a student might have a preference for either a tenure track/tenured or a contingent faculty member, and conditional on deciding to take a given class, takes the class in the quarter in which the student’s preferred faculty member’s status occurs. (This could happen either across terms or within a term if the course is taught simultaneously by a tenure track/tenured faculty member and a contingent faculty member.) This specification helps to rule out, by construction, the possibility that our results are driven by endogenous selection into sections of a given course. In 35.2% of classes taken by first-term freshmen, the student had no choice about the faculty status of an instructor. The results are quite similar whether or not we make the restriction to look only at students with no choice, as seen in the third row of table 2.¹⁰

The available evidence presented thus far all suggests that there is no systematic sorting of students into classes taught by tenure track/tenured versus contingent faculty, but we can also investigate this question directly by estimating sorting regressions in which we regress student attributes against a contingent faculty dummy variable and include fixed effects for the class-term-year combination. We find no evidence of differential sorting by student characteristics in the admissions data. In this model, contingent faculty attract students with a 0.0002 higher academic pre-

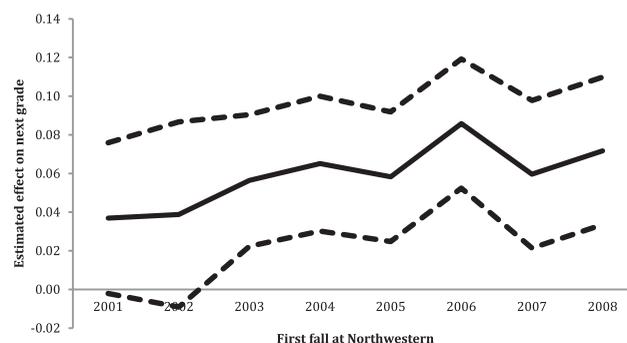
¹⁰ We can also limit ourselves to the small number of cases—17% of students, 15% of student-class observations—in which a student indicates no intended major preference at the time of entry to Northwestern. The results for this very restricted group are similar to those reported in the table. In models with student fixed effects, the coefficient on contingent faculty is 0.120 (0.138 for students with no choice) when the dependent variable is the probability of taking another class in the subject and 0.089 (0.088 for students with no choice) when the dependent variable is the grade in the next class taken in the subject. All of these coefficient estimates are statistically significant at the 5% level or better. However, we do not have sufficient power to estimate our preferred specification—with both student fixed effects and next-class fixed effects—with the restricted set of students who have no intended major at the time of entry.

paration indicator (standard error of 0.0078) and with 0.886 points lower SAT scores or SAT equivalent (standard error of 1.259). Moreover, there is no evidence of differential sorting by a broader set of student characteristics observed in Northwestern's administration of the freshman survey of the Consortium on Financing Higher Education (COFHE): contingent faculty draw students whose mothers are 0.5 percentage points less likely to be college graduates (standard error of 0.7), who are 0.01 percentage points more likely to have named Northwestern as their first-choice school (standard error of 1.04), who are 0.07 percentage points less likely to have parents who are living together (standard error of 0.08), who are 0.1 percentage points more likely to be native English speakers (standard error of 0.5), who are 0.01 percentage points more likely to live within 10 miles of Northwestern (standard error of 0.35), and who are 1.0 percentage points more likely to be female (standard error of 1.4). In summary, we could find no evidence that students differentially self-select into taking courses by tenure track/tenured or contingent faculty along any observable dimension.

We next move to our preferred model specification—one with both student fixed effects and next-class fixed effects—reported in the fourth row of table 2. In this model we cannot study the relationship between contingent status and the likelihood of taking another class in the subject because by default, all students have taken another class in the subject. Moreover, we cannot limit ourselves to students with no choice because we must compare those who took the introductory class in subject A with a contingent faculty member to those who took that same class with a tenure line professor to have variation when we control for next-class fixed effects. When we estimate this highly parameterized model, we still find that having an initial experience in a subject with a contingent faculty member increases a student's performance in subsequent experiences with the subject. The point estimates are around half the size of those found in the specification that includes only student fixed effects but are still statistically significant and sizable in magnitude, especially given that the typical student's grade in the next class is a robust 3.39 out of 4.

Because Northwestern relies somewhat more on contingent faculty today than it did a decade ago¹¹ and because Northwestern freshman classes have become progressively more qualified over time, we also investigate whether the estimated impact of having a contingent faculty member is trending over time. (Of course, we have already ruled out the primary effects of temporal trends by including both student fixed effects and next-class fixed effects.) As can be seen in figure 1, when we estimate our highly parameterized model year by year, we still observe a positive relationship between having a contingent faculty member and subsequent grades in the subject in every year. In addition, there

FIGURE 1.—ESTIMATED EFFECTS OF HAVING A CONTINGENT FACULTY MEMBER ON SUBSEQUENT PERFORMANCE IN THE SUBJECT, BY FRESHMAN YEAR COHORT



These estimates are computed for the model with student fixed effects and next-course fixed effects. Data include all students who enrolled in their first quarter at Northwestern University during the fall terms from 2001 to 2008. The solid line shows the estimated next-grade effect, with the dashed lines indicating a 90% confidence interval.

is no evidence of a meaningful temporal pattern in these results, suggesting that any trends over time in the use or utility of contingent versus tenure line faculty members is not driving the findings that we report.¹²

IV. Differences by Faculty Member Characteristics

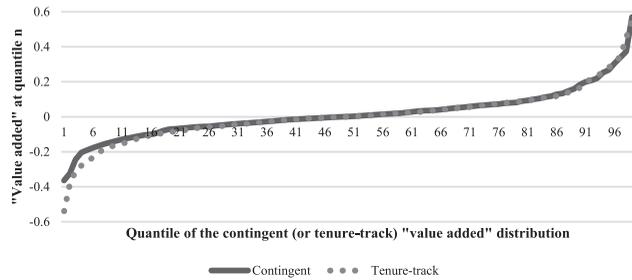
Our results suggest that on average, first-term freshmen learn more from contingent faculty members than they do from tenure track/tenured faculty. But are these differences constant across the entire spectrum, or is it the case that most tenure track/tenured faculty members perform similarly to most contingent faculty members and the differences are due either to the best contingent faculty teachers substantially outperforming the best tenure track/tenured teachers or to the worst tenure track/tenured teachers performing considerably worse than the worst contingent faculty teachers? And if the difference is not constant across all faculty members, are our results driven by a handful of outliers or a by larger swath of the distribution?

To explore this question, in figure 2 we compare the distributions of value added of individual contingent faculty teachers and tenure track/tenured teachers, in which we plot a variant of the cumulative density function where the percentile in each distribution is on the horizontal axis and the corresponding value-added measure is on the vertical axis. (We choose to present the CDF in this manner because it makes clear exactly where in the distribution our results are coming from.) An individual instructor's value added is an instructor-specific fixed effect retained from our preferred specification (in which we estimate instructors' effect on grade points earned in the next course, controlling for both

¹¹ The typical freshman in fall 2001 took 38.9 percent of first-term courses from contingent faculty, as compared with 41.6 percent for the typical freshman in fall 2008.

¹² One might also be concerned that changing grading standards over time are potentially driving our results. However, grading standards have remained quite flat at Northwestern during this time period. While the average next-course grade did rise modestly from 3.33 for fall 2001 entrants to 3.41 for fall 2008 entrants, student qualifications also rose during this time period, with SATs increasing from 1376 for fall 2001 entrants to 1421 for fall 2008 entrants, so that average qualifications-adjusted grades actually fell slightly over our time horizon. We describe in section V our method for adjusting grades for student qualifications.

FIGURE 2.—DISTRIBUTIONS OF ESTIMATED VALUE ADDED: CONTINGENT VERSUS TENURE-TRACK FACULTY



These estimates of instructor grade point effects are computed for the model with student fixed effects and next-course fixed effects. Data include all students who enrolled in their first quarter at Northwestern University during the fall terms from 2001 to 2008. This figure plots the distribution of instructor fixed effects, by faculty member type.

student fixed effects and next-course fixed effects). As can be seen in the figure, the top three-quarters of the contingent faculty and tenure track/tenured faculty distributions are virtually perfectly overlapping, so the most outstanding contingent faculty members and most outstanding tenure track/tenured faculty members perform essentially identically, and the same is true at other points in the distribution such as the median. But the bottom quarter of the tenure track/tenured faculty have lower value added than the bottom quarter of the contingent faculty, and this difference is substantial for the bottom 13% of the distribution (around the weakest 150 instructors, by our definition). It is clear that our results are not being driven by a handful of outliers, but it is also clear that the difference in average outcomes is due to the differences at the bottom of the value-added distribution.

In some ways, this is exactly what we might have expected: contingent faculty members who are hired to teach and who perform relatively poorly are less likely to be renewed than are those who perform well, while tenure track faculty who are relatively poor teachers may be promoted and retained for reasons other than their teaching ability. That said, the presence of these differences begs the question of whether certain differences between contingent faculty members and tenure track/tenured faculty members can explain our findings. There are no administrative records that can address this question directly, so we collected curriculum vitae available through extensive web searching in order to measure some differential attributes of contingent and tenure track/tenured faculty.¹³ Two measures that are the most directly observable are years of experience (calculated based on time since Ph.D. and employment history) and native language (calculated based on the country in which a faculty member earned his or her bachelor's degree or its equivalent).¹⁴ Tenure track/tenured

¹³ We found sufficiently complete curriculum vitae on the web for 87.1% of faculty members.

¹⁴ We consider a faculty member to be a native English speaker if he or she attended an undergraduate institution in the United States, Canada, Great Britain, Ireland, Australia, New Zealand, or South Africa. All results are comparable if we look exclusively at those educated as undergraduates in the United States or some subset of these countries.

TABLE 3.—ESTIMATED EFFECTS OF HAVING A CONTINGENT FACULTY MEMBER ON SUBSEQUENT COURSE PERFORMANCE IN THE SUBJECT, BY INSTRUCTOR EXPERIENCE: MODELS WITH STUDENT FIXED EFFECTS AND NEXT-CLASS FIXED EFFECTS

Instructor Experience	All Classes	Classes Outside Declared Major
5 years or fewer	0.042 (0.031)	0.048 (0.033)
6–12 years	0.110*** (0.036)	0.140*** (0.040)
13 years or more	0.103*** (0.019)	0.108*** (0.018)

Data include all students who enrolled in their first quarter at Northwestern University during the fall terms between 2001 and 2008. Each cell represents a different model specification. Standard errors adjusted for clustering at the instructor level are reported in parentheses beneath coefficient estimates. Next-class information is recorded for the first time a student takes a second class in a given subject area. Intended majors are recorded by the office of admissions. Coefficients are statistically distinct at ***1%, **5%, *10%. Data come from 15,662 students taking 56,599 first-quarter classes.

faculty are modestly more likely to have attended undergraduate institutions in English-speaking countries (86.3% versus 79.2% for contingent faculty, $p = 0.123$) and average dramatically more experience (21.9 years versus 11.6 years for contingent faculty, $p = 0.000$). Given these differences, we consider two models in which we control for the country of undergraduate education (which we call “home country”) and years of experience. In the fifth row, of table 2, we include years of experience as a linear control, and in the sixth row, we categorize years of experience into six groups.¹⁵ While the magnitudes of coefficients are modestly lower (the point estimates are between 16% and 35% smaller) in models in which we condition on home country and years of experience, we continue to find reasonably large and consistently robust evidence that on average, students learn more in first-term classes taught by contingent faculty members than they do when these classes are taught by tenure track/tenured faculty members.

The (modest) differences that do exist between the results of models that do not control for experience and home country and those that do are driven by the control for experience levels. Therefore, in table 3 we stratify faculty members into three groups based on measured experience: those with five or fewer years of experience, those with six to twelve years of experience, and those with thirteen or more years of experience.¹⁶ We observe that while students apparently learn more on average from contingent faculty members than from tenure track/tenured faculty in first-term courses, the results are concentrated in the two higher-experience groups rather than in the low-experience group.¹⁷

Faculty members may perform differently depending on their status at the university. We therefore divide tenure line

¹⁵ We also include missing data flags for the cases in which we are missing either experience levels or home country. Our results are very similar regardless of how we treat faculty members with missing data.

¹⁶ This is as close as we could come to constructing thirds of the contingent faculty experience distribution.

¹⁷ The reported results do not control for home country, but controlling for home country barely changes the results. For example, the estimates in the first column of table 3 would be 0.033 (standard error = 0.037), 0.100 (standard error = 0.037), and 0.104 (standard error = 0.018) were we to control for home country.

TABLE 4.—ESTIMATED EFFECTS OF INSTRUCTOR TYPE ON SUBSEQUENT COURSE PERFORMANCE IN THE SUBJECT, BY OTHER INSTRUCTOR ATTRIBUTES, MODELS WITH STUDENT FIXED EFFECTS, AND NEXT-CLASS FIXED EFFECTS

Faculty Type ^a	All Classes	Classes Outside Intended Major
Untenured professor (29.7% of tenure-track)	0.005 (0.010)	-0.017 (0.011)
Part-time contingent faculty (23.1% of contingent)	0.052*** (0.012)	0.060*** (0.013)
Full-time contingent faculty (76.9% of contingent)	0.062*** (0.009)	0.081*** (0.010)

Data include all students who enrolled in their first quarter at Northwestern University during the fall terms between 2001 and 2008. Each cell represents a different model specification. Standard errors adjusted for clustering at the instructor level are reported in parentheses beneath coefficient estimates. Next-class information is recorded for the first time a student takes a second class in a given subject area. Intended majors are recorded by the office of admissions. Full-time contingent faculty members teach at least four courses per year at Northwestern. Coefficients are statistically distinct at ***1%, **5%, *10%. Data come from 15,662 students taking 56,599 first-quarter classes.

^aComparison group is tenured professor.

faculty into tenured professors versus untenured professors (at the time of the course), and we divide contingent faculty into full time (measured as teaching four or more courses per year at Northwestern) and part time (measured as teaching three or fewer courses per year).¹⁸ We found that 29.7% of tenure track/tenured faculty in first-term courses are untenured, while 76.9% of contingent faculty teaching first-term courses are full time.¹⁹ Table 4 presents the estimated effects of having a faculty member in one of these mutually exclusive groups—untenured professors, part-time contingent faculty, and full-time contingent faculty—with each group’s estimated effect compared to tenured faculty.

As can be seen in the table, tenure track professors without tenure have teaching outcomes that are on average equal to those of tenured professors, while both full-time and part-time contingent faculty members apparently outperform tenured professors in the first-term classroom. The estimated effect of full-time contingent faculty is modestly higher than that for part-time contingent faculty, but the differences between these two groups are not statistically different from 0. It is important to note that the overwhelming majority of part-time contingent faculty at Northwestern still have long-term relationships with the university, so we do not equate part-time faculty at Northwestern with “one-off” adjunct instructors; rather, a large fraction of part-time contingent faculty members teach a course or two at Northwestern in addition to their regular professional careers. This may be the reason that we find a positive effect of part-time contingent faculty at Northwestern, while previous research, such as Bettinger and Long (2010), found less favorable effects: part-time instructors at Northwestern almost always have long-term relationships with the university, while adjuncts hired to fill temporary vacancies may not have the same sense of commitment to an institution and its students.

We also sought to observe whether outstanding researchers are better or worse introductory teachers than are less

¹⁸ We have also estimated models in which we consider a three-course teacher as a full-time contingent faculty member, and the results are very similar.

¹⁹ A larger fraction of arts and sciences contingent faculty are full time.

distinguished researchers. Northwestern has, since 1988, annually recognized a set of scholars who have received extraordinary honors for their scholarship.²⁰ Around 40% of tenured faculty members at Northwestern have been recognized at least once over the past 25 years as an extraordinary researcher. When we treat these faculty members as a separate group, we find no difference in teaching outcomes compared to tenured faculty who have not received the recognition, and this does not significantly affect the other results in table 4.²¹

V. Differences by Subject and Student Qualifications

Are the results the same for all students and all subjects, or are they present in some cases but not in others? In order to investigate these questions, we next divide the course subjects along two dimensions. First, we split the subjects into thirds based on the SAT scores of incoming students who intend to major in that discipline; we interpret this as a measure of the perceived challenge of a subject by incoming students. Second, we split the subjects into thirds based on a measure of the grading standards of faculty teaching that subject. We calculate grading standards by regressing grades against observed student qualifications;²² we call the departments that award higher-than-predicted grades “higher-grading subjects.”²³ These two measures are negatively correlated: the correlation between the average SAT scores of intended majors in a department and the grades that the department awards is -0.69 . There is enough of a discordance between the two to make reporting both measures meaningful. For instance, though the highest-grading subjects generally fall into the low-SAT subject group, 33.3% of the subjects with the highest grades are in the middle SAT group, and 4.1% are in the highest SAT group.

We report the results of these splits for our preferred model specification—with student fixed effects and next-class fixed effects—in table 5. As can be seen, the estimated effects of contingent instructors for a first-term course

²⁰ Reasons for being honored by Northwestern include recognition by the leading scholarly organization in their fields, receipt of prestigious fellowships such as MacArthur or Guggenheim Fellowships, election to the American Academy of Arts and Sciences, and comparable achievements. We use this as a measure of research productivity because other potential measures, such as the value of grants or number of citations, would likely not have similar meaning across academic subjects.

²¹ Tenured faculty who were recognized between 1988 and 2013 for research have an effect size of -0.001 (standard error = 0.011) relative to other tenured professors (estimate = 0.006, standard error = 0.011 for classes outside of intended major). Treating only tenured faculty who have not received research recognition as the comparison group for table 4 instead of all tenured faculty does not qualitatively affect the results reported: point estimates for contingent faculty types remain positive and significant at the 1% level, and the magnitudes are within 7% of the reported estimates, and untenured tenure-track faculty remain similar on average to tenured faculty not recognized for research.

²² Betts and Grogger (2003) and Figlio and Lucas (2004) measure grading standards in similar ways by comparing grades awarded to some external benchmark of predicted grades.

²³ We are restricted by the registrar from identifying specific departments in this paper.

TABLE 5.—ESTIMATED EFFECTS OF HAVING A CONTINGENT FACULTY MEMBER ON SUBSEQUENT PERFORMANCE IN THE SUBJECT: DIFFERENTIAL EFFECTS BY MAJOR “CHALLENGE”: STUDENT FIXED EFFECTS AND NEXT-CLASS FIXED EFFECTS

	All Classes	Classes Outside Declared Major
Subjects divided by SAT score of students (divided into thirds)		
Highest SAT subjects	0.084*** (0.013)	0.099*** (0.013)
Middle SAT subjects	0.063*** (0.017)	0.091*** (0.016)
Lowest SAT subjects	0.030*** (0.014)	0.043*** (0.016)
Subjects divided by typical grade (divided into thirds)		
Lowest-grading subjects	0.077*** (0.013)	0.097*** (0.013)
Middle-grading subjects	0.056*** (0.015)	0.071*** (0.015)
Highest-grading subjects	0.041*** (0.015)	0.061*** (0.018)

Data include all students who enrolled in their first quarter at Northwestern University during the fall terms between 2001 and 2008. Each cell represents a different model specification. Standard errors adjusted for clustering at the instructor level are reported in parentheses beneath coefficient estimates. Next-class information is recorded for the first time a student takes a second class in a given subject area. Intended majors are recorded by the office of admissions. Coefficients are statistically distinct ***1%, **5%, and *10%. Data come from 15,662 students taking 56,599 first-quarter classes. Grading levels of subjects are determined by comparing average residuals of a regression of grades on observed student characteristics; results are qualitatively unchanged if grading levels are unadjusted.

are positive for all sets of subjects, regardless of grading standards or perceived challenge. That said, the estimated effects are strongest for the subjects with tougher grading standards (the relatively low-grading classes) and for those that attract the most qualified students. The estimated effect of having a contingent faculty member on subsequent grades is more than twice as large in the high-SAT subjects as in the low-SAT subjects, and is also substantial when comparing the hardest-grading to the easiest-grading subjects.²⁴

This pattern of results could be due to the effects of being strongest for these groups of classes or because of ceiling effects: perhaps in the easiest-grading subjects, most students earn top grades and there is little opportunity for distinction. As can be seen in table 6, it is not the case that everyone earns a top grade, even in the relatively easy-grading subjects. This table presents the percentage of students earning a grade of A– or A in each group of subjects, broken down by the student’s academic indicator (the admission office’s preenrollment prediction of a student’s academic success at the university). We see that in the easiest-grading third of subjects, 13% of students with the highest academic preparation earn grades of B+ or lower compared with 28% of students with relatively low aca-

²⁴ We have also divided the departments into those that are exceptionally strong in research versus their competitors elsewhere and those that are not. In 74.2% of cases, students take classes ranked in the top twenty nationally in the most recent National Research Council rankings. If we restrict our analyses to the 25.8% of classes in subjects ranked outside the top twenty nationally, we continue to observe positive and significant estimated effects of contingent faculty in our model with student and next-course fixed effects. The coefficient on contingent faculty is 0.049 (standard error of 0.016) for all students and 0.067 (standard error of 0.017) for students taking courses outside their intended major.

TABLE 6.—PERCENTAGE OF STUDENTS EARNING A– OR BETTER IN A CLASS, BY SUBJECT AND ACADEMIC INDICATOR

	Highest Academic Preparation	Middle Academic Preparation	Relatively Low Academic Preparation
All classes	67.2%	56.3%	46.3%
Subjects divided by the average SAT scores of freshmen with intended majors (divided into thirds)			
Highest SAT subjects	56.0%	38.1%	23.6%
Middle SAT subjects	76.2%	64.0%	47.6%
Lowest SAT subjects	82.7%	73.1%	61.5%
Subjects divided by typical grade in classes (divided into thirds)			
Lowest-grading subjects	54.5%	36.5%	21.8%
Middle-grading subjects	71.4%	57.5%	40.4%
Highest-grading subjects	87.4%	80.8%	71.5%

ademic preparation.²⁵ While there is certainly more room for grade dispersion in the toughest-grading subjects, where 45% of students with the highest preparation earn a B+ or lower and 78% of students with relatively low academic preparation do the same, the point is that in no subject and for no group of students is a grade of A or A– a forgone conclusion. Nonetheless, while we cannot say for certain whether the stronger results for harder-grading subjects and those attracting higher-rated students are due to the effects of faculty status truly being greater for those subjects or whether there is simply more room for grade dispersion in those subjects, the key finding is that we observe advantages for contingent faculty across all subjects, regardless of our measure of perceived challenge or grading standards.

In table 7 we split the population of students by student academic preparation and then split subjects by both SAT scores of incoming intended majors and by our measure of faculty grading standards. The rationale for doing this is that students who are relatively less well prepared academically are more representative of the college-going population in the United States as a whole and can therefore assist to some degree in assessing the potential external relevance of our findings. While we find that the best-prepared students at Northwestern appear to perform about the same regardless of whether their first class in the subject was taught by a contingent faculty member or a tenure track/tenured professor, the estimated positive effect of having a contingent faculty member is present and strongly statistically significant for all other groups of students. Moreover, there appears to be an interaction between class type and student qualifications. While there is no apparent relationship between instructor type and student outcomes for the top-rated students, a clear pattern emerges for the other two groups of students. For students with middle academic preparation, the estimated relationship between instructor type and subsequent outcomes is about the same for the subjects attracting relatively low- and midlevel students but substan-

²⁵ It is important to recall that even the relatively marginal students at Northwestern are still very highly qualified. The average SAT (or ACT equivalent) among students with relatively low academic preparation is still a robust 1316.

TABLE 7.—ESTIMATED EFFECTS OF HAVING A CONTINGENT FACULTY MEMBER ON SUBSEQUENT PERFORMANCE IN THE SUBJECT: DIFFERENTIAL EFFECTS BY STUDENT ACADEMIC PREPARATION: STUDENT FIXED EFFECTS AND NEXT-CLASS FIXED EFFECTS

	Highest Academic Preparation	Middle Academic Preparation	Relatively Low Academic Preparation
All classes	0.028 (0.021)	0.062*** (0.011)	0.058*** (0.022)
Classes outside declared major	0.024 (0.023)	0.073*** (0.011)	0.098*** (0.026)
Subjects divided by the average SAT scores of freshmen with intended majors (divided into thirds)			
Highest SAT subjects	0.013 (0.026)	0.099*** (0.018)	0.168*** (0.039)
Middle SAT subjects	0.055 (0.072)	0.059*** (0.022)	0.125*** (0.048)
Lowest SAT subjects	0.029 (0.037)	0.052*** (0.019)	0.004 (0.031)
Subjects divided by typical grade in classes (divided into thirds)			
Lowest-grading subjects	0.007 (0.026)	0.094*** (0.018)	0.175*** (0.038)
Middle-grading subjects	0.075 (0.056)	0.058*** (0.020)	0.069* (0.039)
Highest-grading subjects	0.013 (0.050)	0.059*** (0.023)	0.027 (0.040)

Data include all students who enrolled in their first quarter at Northwestern University during the fall terms between 2001 and 2008. Each cell represents a different model specification. Standard errors adjusted for clustering at the instructor level are reported in parentheses beneath coefficient estimates. Next-class information is recorded for the first time a student takes a second class in a given subject area. Intended majors are recorded by the office of admissions. Coefficients are statistically distinct at ***1%, **5%, *10%. Data come from 15,662 students taking 56,599 first-quarter classes.

tially larger for the subjects attracting the most qualified students and about the same for easiest and middle-grading subjects but considerably larger for the toughest-grading subjects. For students with relatively low academic preparation, the monotonic relationships are even more pronounced, with by far the strongest estimated results of all observed for the relatively marginal students at Northwestern taking the toughest-grading subjects and those attracting the most qualified students. Note from table 6 that the gap in the percentage receiving an A or A– between the toughest-grading and the easiest-grading subjects is 49.7 percentage points for those with relatively low academic preparation, but nearly as high (44.3 percentage points) for those with middle academic preparation and still quite high (32.9 percentage points) for those with the highest academic preparation. Therefore, this pattern of estimated effects of instructor type broken down by student qualifications and subject type suggests that these findings are likely due to genuinely differential effects of instructor type across subject and student preparation rather than just pure ceiling effects or the differential likelihood of earning higher grades in some subjects versus others.

In sum, we found that the strong and significant effect of contingent faculty on our measure of deep learning held for all subjects, regardless of grading standards or the qualifications of the students the subjects attracted. The apparent benefits of taking classes from contingent faculty were particularly strong for tougher-grading subjects and those that attracted the most qualified students, and the benefits were enjoyed more by the less academically qualified students than by the more academically qualified students—the big-

gest gains to taking courses from contingent faculty were for relatively weak students taking courses in the toughest-grading subjects.

VI. Conclusion

Our findings suggest that contingent faculty at Northwestern not only induce first-term students to take more classes in a given subject than do tenure line professors, but also lead the students to do better in subsequent course work than do their tenure track/tenured colleagues. This result is driven by the fact that those in the bottom quarter of tenure track/tenured faculty (as indicated by our measure of teaching effectiveness) have lower “value added” than their contingent counterparts.

How generalizable are these results? Because a key part of our identification strategy is to limit our analysis to first-term freshman undergraduates, the evidence that contingent faculty produce better outcomes may not apply to more advanced courses. Further, Northwestern University is among the most selective and highly ranked research universities in the world, and its ability to attract first-class contingent faculty may be different from that of most other institutions. Importantly, a substantial majority of contingent faculty at Northwestern are full-time faculty members with long-term contracts and benefits, and therefore may have a stronger commitment to the institution than some of their contingent counterparts at other institutions. Northwestern’s tenure track/tenured faculty members may also have different classroom skills from those at other schools. Finally, Northwestern students come from a rarefied portion of the preparation distribution and are far from reflective of the general student population in the United States. That said, our results are strongest for the students and subjects that are most likely to generalize to a considerably wider range of institutions: the benefits of taking courses with contingent faculty appear to be stronger for the relatively marginal students at Northwestern (although they are still very well-prepared students), and our results are similar for top-ranked departments as for lower-ranked departments.

Our identification strategy and setting may help to explain why our results find a more positive effect from contingent faculty members than the earlier literature does. Because contingent faculty members at Northwestern tend to have considerably different contracts than do contingent faculty members at many other institutions, our results may be better thought of as the effects of taking classes with designated teachers, albeit a group of designated teachers who can be fired in the event of poor teaching, rather than generalized results about contingent faculty. Our work also differs from the prior literature in that we look not only at the likelihood that a student takes a subsequent class, but also at the likelihood of success in that class once the student enrolls. It is therefore somewhat difficult to compare our results to previous findings because our treatment of interest is considerably different, as is our outcome of inter-

est: a measure of the effects of faculty of different types on students' deep learning. In addition, our identification strategy differs from the closest paper in the literature to our own, Bettinger and Long (2010), in ways that might partially explain the differences. Because of their identification strategy, Bettinger and Long are better able to identify the effects of short-term transient adjuncts, while at Northwestern University, most of the contingent faculty are full time and even the part-time contingent faculty members generally have long-term relationships with the university.

There are many aspects relating to changes in the tenure status of faculty, from the impact on research productivity to the protection of academic freedom. But certainly learning outcomes are an important consideration in evaluating whether the observed trend away from tenure track/tenured toward contingent faculty is good or bad. Our results provide evidence that the rise of full-time designated teachers at U.S. colleges and universities may be less of a cause for alarm than many assume.

REFERENCES

- Bettinger, Eric P., and Bridget T. Long, "The Increasing Use of Adjunct Instructors at Public Institutions: Are We Hurting Students?" (pp. 51–69), in Ronald G. Ehrenberg, ed., *What's Happening to Public Higher Education? The Shifting Financial Burden* (Baltimore: Johns Hopkins University Press, 2006).
- "Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes," this REVIEW 92, no. 3 (2010), 598–613.
- Betts, Julian R., and Jeff Grogger, "The Impact of Grading Standards on Student Achievement, Educational Attainment, and Entry-Level Earnings," *Economics of Education Review* 22 (2003), 343–352.
- Carrell, Scott E., and James E. West, "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy* 118 (2010), 409–432.
- Ehrenberg, Ronald G., "American Higher Education in Transition," *Journal of Economic Perspectives* 26 (2012), 193–216.
- Ehrenberg, Ronald G., and Liang Zhang, "Do Tenured and Tenure Track Faculty Matter?" *Journal of Human Resources* 40 (2005), 647–659.
- Figlio, David N., and Maurice E. Lucas, "Do High Grading Standards Affect Student Performance?" *Journal of Public Economics* 88 (2004), 1815–1834.
- Figlio, David N., Morton O. Schapiro, and Kevin B. Soter, "Are Tenure Track Professors Better Teachers?" NBER working paper 19406 (September 2013).
- Hoffmann, Florian, and Philip Oreopoulos, "Professor Qualities and Student Achievement," this REVIEW 91, no. 1 (2009), 83–92.
- June, Audrey Williams, "Adjuncts Build Strength in Numbers: The New Majority Generates a Shift in Academic Culture," *Chronicle of Higher Education*, November 5, 2012.
- McPherson, Michael S., and Morton Owen Schapiro, "Tenure Issues in Higher Education," *Journal of Economic Perspectives* 13 (1999), 85–98.
- Wilson, Robin, "Tenure, RIP: What the Vanishing Status Means for the Future of Education," *Chronicle of Higher Education*, July 4, 2010.